R842-09   ERIC REPORT RESUME

ED 010 162   1-30-67   24   (REV)
EFFECTS OF REDUCED LOADS AND INTENSIVE INSERVICE TRAINING UPON THE
CLASSROOM BEHAVIOR OF BEGINNING ELEMENTARY TEACHERS.
HITE, F. HERBERT * AND OTHERS
XL079510 OFFICE OF SUPT. OF PUBLIC INSTRUC., OLYMPIA, WASH.
CRP-2973
XL079700 WASHINGTON STATE UNIV., PULLMAN
BR-5-0360
  -   -66
EDRS PRICE MF-$0.27 HC-$5.64   141P.

INSERVICE TEACHING,
*INSERVICE TEACHER EDUCATION +*TEACHER EVALUATION,
ELEMENTARY SCHOOL TEACHERS, *WORK EXPERIENCE PROGRAMS,
*BEGINNING TEACHERS, TEACHER ATTITUDES, EXPERIMENTAL PROGRAMS,
*EFFECTIVE TEACHING, BEHAVIOR DEVELOPMENT, PULLMAN, WASHINGTON

A STUDY WAS MADE TO SEE WHETHER EXPERIMENTAL TREATMENTS INVOLVING REDUCED WORK LOADS AND INTENSIVE INSERVICE INSTRUCTION WOULD AFFECT THE PERFORMANCES AND ATTITUDES OF BEGINNING TEACHERS. IT WAS HOPED THAT THE STUDY DATA WOULD SHOW TO WHAT EXTENT AN INTERNSHIP PROGRAM FOR CAREER TEACHERS WOULD BE JUSTIFIED. THE EXPERIENCE IN THIS STUDY WITH THE APPRAISAL TECHNIQUES USED TO EVALUATE BEGINNING TEACHERS' PERFORMANCES SUGGESTS THAT EXPERIENCED TEACHERS CAN BE TRAINED TO USE AN OBSERVATION CHECKLIST AND ARRIVE AT A FAIR DEGREE OF AGREEMENT WITH OTHER SUCH TRAINED OBSERVERS. (LP)

5-0360

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

EFFECTS OF REDUCED LOADS AND INTENSIVE INSERVICE TREATMENT UPON THE CLASSROOM BEHAVIOR OF BEGINNING ELEMENTARY TEACHERS

Cooperative Research Project No. 2973 (Bureau No. 5-0360)

F. Herbert Hite, Principal Investigator

Washington State Superintendent of Public Instruction, and
Washington State University, Pullman

Harry L. Garrison, Seattle Schools
G. Paul Killian, Washington State University

1966

ED010162

The research reported herein was supported by the Cooperative Research Program of the Office of Education, U. S. Department of Health, Education, and Welfare.

ACKNOWLEDGEMENTS

The principal investigator here expresses his appreciation to his associates, Harry L. Garrison, of the administrative staff of the Seattle Public Schools, and G. Paul Killian, Washington State University, for their assistance in the conduct of this investigation.

Over 400 school, university and state department workers participated in the project. The staff especially acknowledges the financial and moral support of the school districts of Seattle, Shoreline, Edmonds, Bellevue and Lake Washington. Their contributions to this project were essential.

State Superintendent of School, Louis Bruno, and Wendell C. Allen, Ray Jongeward and Alan Metcalf of his staff materially assisted the project. The research office of the State Superintendent provided the computer analysis of the project data.

Finally the principal investigator expresses his appreciation to the following persons who served as an advisory committee to this project:

Homer Boroughs, Jr., University of Washington
Harry L. Garrison, Seattle Schools
Edward D. Gibbs, University of Puget Sound, Tacoma
Margaret Heller, Principal, Vancouver, Washington Schools
Alexander H. Howard, Central Washington State College, Ellensburg
Joseph P. Lassoie, Tacoma Schools
Roland B. Lewis, Eastern Washington State College, Cheney
John Porter, Edmonds Schools
Elmer B. Siebrecht, Seattle Pacific College
William C. Sorenson, Spokane Schools
Ralph H. Thompson, Western Washington State College, Bellingham
Earl M. Towner, Seattle Schools

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS ii

LIST OF TABLES iv

Chapter

I. STATEMENT OF THE PROBLEM 1

II. OBJECTIVES AND HYPOTHESES OF THE STUDY 4

III. RELATED LITERATURE 6

IV. PROCEDURES 13

V. ANALYSIS AND DISCUSSION OF THE DATA 35

VI. SUMMARY, CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH 75

BIBLIOGRAPHY 82

APPENDIX 84

A. FORMS USED IN THE STUDY 85

B. SUMMARY TABLES 112

LIST OF TABLES AND ILLUSTRATION.

Figure

Page

1. Mean Ratings on all Ten Criteria 59

Table

1. Schedule and Rotation of Teachers and Observers in Teacher . . Appraisal Study. 17
2. Observer Correlations in the Pilot Study for the Stanford Form 22
3. Observer Correlations in the Pilot Study for the Ryans Form. . 22
4. Product Moment Correlations of Pilot Study Subjects' Grade Point with Observer Ratings of their Teaching. 25
5. Mean Total Scores on Teaching Performance by Pilot Study Subjects According to College Academic Major 25
6. Teaching Performances of Pilot Study Teachers Compared to Grades in Student Teaching 26
7. Teaching Performances of Pilot Study Subjects Compared to Grade Level Assignment 26
8. Observer Correlations by Team by Round by Behavior 37
9. Correlations Obtained When Two Observers' Ratings Were Omitted 39
10. Reliability of Observer Appraisals by Team by Round by Behavior 40
11. Inter-item Correlations of Round 4 Ratings on Each of the Ten Performances, Individually and in Combinations 41
12. Observer Evaluation of Teaching Behaviors. 43
13. Mean Performance Ratings of the Four Treatment Groups on First Round of Observation 45
14. Mean Scores of Four Treatment Groups on Observation Rounds 1 and 2. 47
15. Analysis of Covariance: Significance of Differences of Mean Scores Among the Four Groups, Each of 18 Subjects, Obtained on the Second Round of Observations (Round 1 Observation Ratings Used as Covariable). 48

Table Page

16. Mean Scores of Four Treatment Groups on Observation Rounds 1 and 3 52

17. Analysis of Covariance: Significance of Differences of Mean Socres Among the Four Groups, Each of 18 Subjects, Obtained on the Third Round of Observations (Round 1 Observation Ratings Used as Covariable) 53

18. Mean Scores of Four Treatment Groups on Observation Rounds 1 and 4. 58

19. Analysis of Covariance: Significance of Differences of Mean Scores Among the Four Groups, Each of 18 Subjects, Obtained on the Fourth Round of Observations (Round 1 Observation Ratings Used as Covariable) 61

20. Direction of Attitude Changes by Treatment Groups Shown on Four Administrations of the Semantic Differential Scale. . 67

21. Net Changes in Attitudes From the Beginning to the End of the Study by the Four Treatment Groups. 67

22. Analysis of Variance Among Subjects Tested Within Four Treatment Groups on Four Rounds of the Semantic Differential Attitude Test 68

23. F Ratios From Analysis of Variance Among Treatment Groups on Each of Four Administrations of the Semantic Differential Attitude Test (N = 92 F of 2.72 required for .05 level of significance) 72

24. Correlation Coefficients of All-College Grade Point Averages and Education Course GPA with Final Ratings of Teaching Performances. 73

25. Data From the Trial Form of the Semantic Differential Attitude Scale. 113

26. Summary of Data on Pilot Study Teacher Ratings as Related to Subjects' Characteristics. 120

27. Correlations Between Observer Pairs on All Rounds 122

28. Summary of Data on the Semantic Differential Attitude Scale by Treatment Groups and Administrations (N = 96). 126

CHAPTER I

STATEMENT OF THE PROBLEM

The purpose of this study was to see whether differential treatment of beginning teachers during their first year of teaching results in significant changes in either the classroom behavior of these teachers and/or their attitudes towards the profession of teaching.

Unlike other professional neophytes, teachers begin at the top. The teacher in her first day has exactly the same responsibilities as the most able and experienced teacher on the staff. The lack of a formal induction procedure for teachers seems to some researchers to account for the very slow rate of professional growth of teachers. For example, Dan Lortie summarizes the typical pattern of the beginning teachers' orientation in three stages: (a) a year or two of struggles to get through each day without major damage to students or self; (B) a period of attempts to innovate (which usually attract unfavorable attention from administrators); and (c) crystallization into conventional practice.[1] Thus does each wave of new teachers arrive at the same stage of mediocrity.

In the State of Washington the State Board of Education, which has legal responsibility for policies governing teacher education, appointed in 1960 a special committee to recommend changes in standards for the education of teachers. This committee, the Teacher Education Standards Revision Committee, was charged with responsibility of recommending changes in the teacher education program in the State of Washington. After a seriss of discussions, the committee recommended a study of the possible effects of a major change in the treatment of the beginning teacher during the first months of service. It was the committee's opinion that teachers adapt the procedures of other teachers with whom they associate during these first months, and only to a small degree apply the procedures that they have studied in their college teacher education programs. The committee also felt that it was very possible that many teachers form their basic attitudes towards their profession in these first months of service. The committee reasoned that a major change in the treatment of teachers should involve different teaching assignments and much more than the usual amount of inservice help for these beginners. Such an effort could make a greater effect on the behavior of the classroom teacher than any other effort in the entire teacher education program.

---

[1] Lortie, Dan. "Teacher Socialization --The Robinson Crusoe Model" The Real World of the Beginning Teacher. Report of the Nineteenth National TEPS Conference. Washington D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1966. p. 59.

Another consideration was the fact that in the State of Washington the first year of teaching is conceived of as part of the basic education of the teacher. After four years of study in an approved teacher education program, a teacher graduates with the provisional certificate. The standard certificate is granted after two years of successful teaching and an additional year of college study. Although this beginning experiience is thus a part of the actual training program, in fact, most school districts have merely reported the two years of experience as successful without seriously evaluating the performance of the teacher, or making searching recommendations for further training of the teacher.

Both researchers and practitioners strongly urge that the period of beginning teaching should be conceived of as the critical stage in the development and refinement of the technology of classroom teaching. For example, the 1965 Conference of the National Commission on Teacher Education and Professional Standards, in "The Real World of the Beginning Teacher," heard a variety of authorities and participants underline the need for new directions in teacher education which would focus on some form of internship instead of a full teaching assignment for the beginning years.[1]

At about the same time the committee was making the recommendation to the State Board, the James R. Conant report on teacher education was published. One of the major recommendations of the Conant report, recommendation number 11, seemed consistent with the recommendation of the Standards Revision Committee.[2]

The recommendation of the committee, and the basis for this project, was that the State should undertake an experimental study to determine whether or not a substantial reduction in the load of the beginning teacher, together with intensive in-service training, might affect the classroom behavior of these teachers and also affect their attitudes toward their profession.

---

[1] See National Education Association, National Commission on Teacher Education and Professional Standards. The Real World of the Beginning Teacher. Report of the Nineteenth National TEPS Conference. Washington D.C.:

[2] Conant, James B. The Education of the American Teacher. McGraw-Hill Co., 1963.

CHAPTER II

OBJECTIVES AND HYPOTHESES OF THE STUDY

The specific objectives which guided this research were:

1. To determine whether or not reduced loads of beginning teachers would affect the classroom behavior of these teachers.

2. To determine whether or not reduced loads would affect the beginning teachers' attitudes towards teaching.

2. To determine whether or not intensive inservice help for the beginning teacher would affect their classroom behavior.

4. To determine whether or not such intensive inservice help would affect the attitudes of these beginning teachers.

Secondary and related objectives of this study were:

1. To determine whether or not personal or profesiional characteristics of beginning teachers modified the effects of reduced loads and intensive inservice training.

2. To arrive at a defensible judgement about the amount and kind of help for beginning teachers necessary to affect these teacher's behaviors.

3. To determine whether or not it would be administratively feasible and experimentally reliable to have trained observers utilize an evaluative check list to appraise the behavior of the beginning classroom teacher.

The following hypotheses are stated in the null form:

1. There will be no significant differences in terms of selected aspects of teachers' classroom behavior between groups of beginning elementary teachers who have a reduction in load compared to other teachers who have no reduced load.

2. There will be no significant differences in terms of their opinions about teaching between groups of beginning teachers who have reduced load and those who have no reduced load.

3. There will be no significant differences in terms of selected aspects of teachers' classroom behavior among groups of be-

ginning teachers who receive different types of inservice instruction during released time periods.

4. There will be no significant differences in terms of their opinions about teaching among groups of beginning teachers who receive different types of inservice instruction during released time periods.

CHAPTER III

[illegible]

Although there has been a great deal of research on teaching effectiveness, th[illegible] answers to the questions of [illegible] major assumption of this study is that it is feasible to evaluate teaching effectiveness by observing a teacher in the act of classroom teaching, and further, that valid instruments exist which are sensitive to changes in teaching behavior as observed by trained educators.

A study which has become a milestone in the research on teacher effectiveness was directed by David G. Ryans. Ryans concluded that:

> Research (is needed) with the refinement of observing and assessing techniques, leading perhaps to the development of behavioral check lists and forced-choice scales to the end of providing more valid behavior-in-process criterion data.[1]

Ryans' research was focused on pupil and teacher classroom behavior. Over [illegible] teachers were observed and rated. Reliability coefficients were determined using paired-observers who made independent judgements. Coefficients from .54 to .86 were obtained. The minimum sample of teachers observed was 43.

Ryans' behavior characteristics included:

Pupil Behavior

Apathetic - ~~Alert~~

Obstructive - Responsive

Uncertain - Confident

Dependent - Initiating

---

[1] David G. Ryans, "Characteristics of Teachers" Washington, D. C.: American Council on Education, 1960, p.400.

CHAPTER III

RELATED LITERATURE

Although there has been a great deal of research on teaching effectiveness, there has not been a great deal of study related to the major assumptions of this project. A major assumption of this study is that it is feasible to evaluate teaching effectiveness by observing a teacher in the act of classroom teaching, and further, that valid instruments exist which are sensitive to changes in teaching behavior as observed by trained educators.

A study which has become a milestone in the research on teacher effectiveness was directed by David G. Ryans. Ryans concluded that:

> Research (is needed) into the refinement of observing
> and assessing techniques, leading perhaps to the development
> of behavioral check lists and forced-choice scales
> to the end of providing more valid behavior-in-process
> criterion data.[1]

Ryans' research was focused on pupil and teacher classroom behavior. Over 3400 teachers were observed and rated. Reliability coefficients were determined using paired-observers who made independent judgements. Coefficients from .54 to .86 were obtained. The minimum sample of teachers observed was 43.

Ryans' behavior characteristics included:

Pupil Behavior

Apathetic - Alert

Obtrusive - Responsive

Uncertain - Confident

Dependent - Initiating

---

[1] David G. Ryans, "Characteristics of Teachers" Washington, D. C.: American Council on Education, 1960, p.400.

Teacher Behavior:

Partial - Fair

Autocratic - Democratic

Aloof - Responsive

Narrow - Broad

A total of four pupil and 18 teacher characteristics were used.

Biddle[1] has proposed a seven variable model for determining teacher effectiveness.

1. Formative experiences of the teacher
2. Teacher properties
3. Teacher behaviors
4. Intermediate effects
5. Long term effects
6. School and community contexts
7. Classroom situations

Approaches such as this have been studied in the past by Barr[2] and McCall[3]. The model variables have a tendency to overlap in areas and, due to the great amount of interaction, this type of model has not been generally accepted by the profession.

Some recent studies in classroom observation have used the approach of studying classroom interaction. Hughes[1] and Flanders[2] have completed research in this approach. Flanders developed an observation instrument with ten categories for observable interaction behaviors. Seven were descriptive of behaviors of teachers, two were of student behaviors, and one was for silence. He distinguished between student and teacher patterns. All categories were observable in terms of direct or indirect influence of the teacher. He compared the patterns with student achievement and attitudes.

---

[1] J. B. Biddle and William J. Ellena, "Contemporary Research on Teacher Effectiveness" Chicago: Holt, Rinehart and Winston, 1964, pp.5-18.

[2] A. S. Barr, "Characteristic Differences in Teaching Performance of Good and Poor Teachers of the Social Studies" Bloomington, Ill.: Public School Publishing Co., 1929.

[3] William A. McCall, "Measurement of Teacher Merit" Raleigh, N. C.: State Department of Education, 1962.

Garrison[3] working with the Stanford Secondary Education Project, developed a teacher appraisal guide based on the Getzel interaction model of social, psychological, and anthropological aspects of a class situation. Four steps of teaching were defined. Twenty-five behavioral statements were incorporated in the four steps; the final statement being an overall appraisal of the teacher. Garrison found significant teacher improvement when the observation appraisal instrument was used as an improvement of instruction device. The attitude of teachers towards observation was improved when used to improve their teaching rather than used as a rating device of the teacher. The reliability coefficients of observer agreement were consistently above .50. He concluded observers could agree on the level of performance and that his instrument was sensitive to levels of performance.

---

[1] Marie Hughes, "Patterns of Effective Teaching, Second Progress Report, Merit Study" Provo, Utah: Provo City Schools, 1961.

[2] Ned Flanders, "Teacher Influence, Pupil Attitudes and Achievement, Studies in Interaction Analysis" Minneapolis: University of Minnesota, U. S. Office of Education Cooperative Research Project No. 397. Mimeographed.

[3] Harry L. Garrison, "Evaluation of Teaching and Learning" unpublished Ed.D. thesis, Stanford University, 1964.

Medley and Mitzel[1] describe other classifications and methods for direct observation and appraisal of classroom interaction. Included in these and the types of criteria investigated are:

Morsh's:[2]

1. Instructor verbal behavior
2. Instructor non-verbal behavior
3. Student behavior

Anderson and Brewer:[3]

1. Domination with evidence of conflict
2. Domination with no evidence of conflict
3. Integration with evidence of working together
4. Nervous habits of students
5. Child domination of other children
6. Non-conformance to teacher demands
7. Social contributions by the child

Mitzel and Medley:[4]

1. Pupil supportive
2. Problem structuring
3. Directive

---

[1] Donald M. Medley and Harold E. Mitzel, "Measuring Classroom Behavior by Systematic Observation," Handbook of Educational Research, ed. N. L. Gage Chicago: Rand McNally and Co., 1963, p. 247-329

[2] J. E. Morsh, Systematic Observation of Instructor Behavior. USAF Personnel Training Research Center Development Report No. AFPTRC-TN-56-52

[3] H. H. Anderson and Helen M. Brewer, "Studies of Teachers' Classroom Personalities II: Effects of Teacher's Dominative and Integrative Contacts on Children's Classroom Behavior," Applied Psychology Monograph, 1945, No. 6

[4] Donald M. Medley and H. E. Mitzel, "Studies of Teacher Behavior: Refinement of Two Techniques for Assessing Teacher's Classroom Behaviors" New York: Board of High Education, City of New York, Division of Teacher Education, Office of Research and Evaluation, 1955.

4. Reproving

Jersild[1] and others:

1. Cooperating activities
2. Experimental activities
3. Critical activities
4. Leadership activities
5. Recitational activities
6. Self-initiated activities
7. Work spirit activities

All of these studies have shown that the classroom observation approach to the appraisal of teaching is feasible. They have asked educators what should take place in a classroom and they have called upon educators to interpret or evaluate what they observe in a classroom.

Another major assumption of this study is that it is both important and possible to assess changes in attitudes of beginning teachers.

Getzels and Jackson refer to the importance of the teacher's attitude as a classroom variable by stating, "The educational impact of an Ichabod Crane or a Mark Hopkins, of a Mr. Chips or a Socrates, is surely not due solely to what he knows, or even to what he does, but in a very real sense to what he is."[2]

Studies of teacher attitudes include one by Callis (1950) who investigated the changes that occur during teacher training and early teaching experience. The measurement instrument used was an extension of the Leeds inventory. The first six months of professional training produced significant changes in the desired direction in 20 per cent of the attitudes (items) while the first six months of experience produced significant changes in 11 per cent of the attitudes(items) in an undesirable direction.[1]

---

[1] A. T. Jersild et al., " An Evaluation of Aspects of the Activity Program in the New York City Public Elementary Schools," Journal of Experimental Education, VIII (April, 1939), 166-207.

[2] J. W. Getzels and P. W. Jackson, "The Teacher's Personality and Characteristics," Handbook of Research on Teaching, ed., N. L. Gage Chicago: Rand McNally & Company, 1963, p. 506.

An additional study of the attitude change of beginning teachers after initial teaching experience was reported by Harry P. Day. The study utilized in part data collected by administering the Minnesota Teacher Attitude Inventory to a group of graduates and to a group of student teachers at Florida State University. This was administered to the two groups before and after an initial teaching experience. The graduates took the test initially upon their return from internship teaching in the public schools. A year later it was administered again to those who had prepared but had not entered teaching. There was a mean loss of 20.0 for those who had taught for one year compared with a mean loss of 1.5 for the non-teaching group. The group which was administered the MATI before and after internship also showed a mean loss of 4.2. In both groups there was a shift toward less desirable attitudes after a teaching experience.[2]

More than 50 other studies using attitude inventories have been summarized by Getzels and Jackson in the "Handbook of Research on Teaching".[3] For this project the staff chose to use the semantic differential technique as reported by Osgood and his associates. As a by-product of their work in experimental semantics; Osgood, Suci and Tannenbaum developed a new approach and rationale for attitude measurement. Attitudes according to Osgood

---

[1] Getzel and Jackson, op. cit., p. 509

[2] Harry P. Day, "Attitude Changes of Beginning Teachers After Initial Teaching Experience," Journal of Teacher Education X, September, 1959, pp. 326-328.

[3] Getzels and Jackson, op. cit., p. 508

"can be ascribed to some basic bipolar continuum with a neutral or zero reference point, implying that they have both direction and intensity and providing a basis for the quantitative indexing of attitudes".[1] The semantic differential technique for measuring meaning is essentially a combination of controlled association and scaling procedures. The subject is provided with a concept to be differentiated and a set of bipolar adjective scales against which to do it. His task is to indicate, for each item, the direction and intensity on a seven-step scale.[2]

Evaluation of teachers' behavior appears likely to be affected by a variety of variables. The research staff assumes that it is particularly important, in measuring teaching behavior, to attempt to control, or compensate for, as many gross variables as possible. In the "Handbook for Research on Teaching", Donald Campbell and Julian Stanley suggest a design which appears to the research staff of this proposal to be applicable to this problem.

> "In Design 4, matching can be recognized as a useful adjunct to randomization but not as a substitute for it: in terms of scores on the pretest or on related variables, the total population available for experimental purposes can be organized into carefully matched pairs of subjects; members of these pairs can then be assigned at random to the experimental or the control conditions. Such matching plus subsequent randomization usually produces an experimental design with greater precision than would randomization alone."[3]

---

[1] Charles E. Osgood, George J. Suci, and Percy H. Tannenbaum, The Measurement of Meaning, Urbana, Illinois: University of Illinois Press, 1957, pp. 189-191

[2] Ibid. p. 20.

[3] Donald T. Campbell and Julian C. Stanley, "Experimental and Quasi-Experimental Designs for Research on Teaching", in Handbook for Research on Teaching, N. L. Gage, Editor, Chicago: Rand, McNally and Co., 1963 P. 219.

CHAPTER IV

PROCEDURE

The Design of the Study

The purpose of the study was to see whether differential treatment of beginning teachers during their first year of teaching results in significant changes in either the classroom behavior of these teachers and/or their attitudes toward the profession of teaching.

A pilot study was carried out in the spring of 1965 with 28 beginning teachers of the Seattle School District. The purpose of this study was to test instruments for assessing attitudes and appraising different classroom behaviors of beginning teachers. The pilot study was also intended to field test procedures for collecting the essential data of the study. The staff felt that it was particularly important to assess procedures for training the observers who would visit beginning teachers to evaluate their behavior. This pilot study was to determine the specific instruments and the procedures to be used in collecting the data for the experiment which was to take place in the fall semester of 1965. An additional purpose of the pilot study was to collect data about the beginning teachers and relate these data to judgements about their classroom performance. The purpose of this effort was to determine which characteristics of teachers might be important ones to use as matching variables in the later experiment.

Approximately 120 beginning elementary teachers were to be the population for the experiment. In order to eliminate many variables which the staff judged not related to the major questions of the study, the teachers selected for the study were to be drawn from as homogeneous a population as practicable. The population was limited to elementary teachers who were women, graduates of teacher-education institutions in the State of Washington of the preceding June, and who were assigned to schools of comparable socio-economic populations. This socio-economic population served by the schools was middle or upper-middle class. After determining the teacher population from which the experimental subjects were to be drawn, two characteristics of beginning teachers, determined by the pilot study, were to be used as matching variables. Matched teachers were then randomly assigned to four different treatment groups.

The four groups were:

1. Group I was to be released of approximately 25 percent of classroom teaching time. This released time was to be used by the teacher for class preparation, and also for conferences with a supervisor who would observe the teacher's performance during the week. In Group I each teacher was visited by a particular supervisor twice a week, and conferred with the superivsor on the basis of this observation twice a week. The supervisors were trained in the techniques involved in using a teacher appraisal guide. This appraisal guide was the same one used to judge performances of all of the teachers in the study.

   The research project funds were used by the school districts participating in the study to hire experienced substitute teachers who took over the classes of Group I teachers for a quarter of each day during the period of the experiment. The same substitute teacher worked with the individual Group I teacher throughout the entire period of the study. In effect, this was a miniature teaching team.

2. Group II teachers were released from approximately 25 per cent of classroom teaching time in the same manner as were Group I teachers. The beginning teachers of Group II used this released time for preparation for their classes, and twice a week visited an experienced teacher with a similar assignment. These experienced teachers, observed by Group II subjects, were selected by the beginning teachers' principals as being teachers expert in some teaching approach applicable to the classroom of the Group II teacher.

3. Group III teachers were given approximately 25 per cent reduction in pupil load as were the first two groups. This reduction in load was accomplished, however, by assigning the Group III teachers to classrooms with approximately 25 per cent fewer pupils than the average assignment of teachers in that school district for that grade level. No other special treatment was given the Group III teacher.

4. Group IV teachers were intended to be a control group. They were to receive no special treatment comparable to that of Groups I, II, or III. They received the orientation provided other beginning teachers not members of

the study, which was the practice in the five cooperating school districts. They were visited by observers and an attitude test was administered them as was the case with the other three groups. It was hoped that the visits by observers would be sufficiently different treatment for Group IV teachers that this would equalize the "Hawthorne Effect" which might be experienced by the other teachers in the study.

. Observers were trained to administer an appraisal check list covering selected criteria for judging the performance of a classroom teacher. These observers were to visit each teacher in teams of three on four occasions. The first round of observations was to take place during the first four weeks of school in the fall. The judgements made on this round of observations were to be used as covariates in analyses of covariance to assess changes in chassroom behavior of all teachers in the study during the experimental period. The observations were repeated at intervals of four and one-half weeks. The last of the four observations was made in January just prior to the conclusion of the fall term. The observer teams were rotated so that each observer worked with any other observer as infrequently as possible, while still visiting any one teacher a minimum number of times. For purposes of scheduling the observations, the total group of teachers were divided into three groups. The schedule for assigning the nine observers to the three groups of teachers for four observation rounds is shown in Table 1

TABLE 1

SCHEDULE AND ROTATION OF TEACHERS AND
OBSERVERS IN TEACHER APPRAISAL STUDY

Teacher Group	Round 1 Obs. Team	Round 2 Obs. Team	Round 3 Obs. Team	Round 4 Obs. Team
1	ABC	ADG	CEF	BHI
2	DEF	BEH	DGI	FAC
3	GHI	CFI	ABH	GDE

On the occasion of each observation, each of the experimental subjects was given a questionnaire to indicate her attitudes towards teaching. Thus there were four administrations of the attitude scale during the study.

The data for the study, then, were the judgements of three independent observers of approximately 120 beginning teachers on four different occasions, and the responses of the beginning teachers to the four different administrations of the attitude scale. These data were to be analyzed by computers to determine whether or not there were significant differences among the groups in terms of either classroom behavior or attitudes.

The Pilot Study

During the spring semester of 1965, 28 beginning teachers in the Seattle School District were visited three times by a team of three trained observers. The observers were trained in the application of two scales for appraising classroom teaching. One was the Ryans'

Classroom Observation Record and the other was the Stanford Teacher Appraisal Competence Guide. Copies of these instruments appear in Appendix A. The two forms were selected because members of the research staff had experience in previous studies with these forms, and also because the two scales seemed to emphasize contrasting criteria. The Ryans Scale seemed to be concerned with the characteristics of the teacher, while the Stanford Guide seemed to be focused upon the teaching act. Previous experience with the Stanford form had been limited to secondary school teachers, and one of the concerns of the staff was whether or not the criteria of this Stanford form were applicable to elementary teaching.

The 28 teachers who were the population for the pilot study were all in their first year of teaching. Approximately half of the group had begun teaching during the same semester as the pilot study. The other members of the group had begun at different times during the preceding semester. They were chosen because they were the least experienced teachers in the Seattle School System. This lack of experience on the part of the pilot study subjects, was considered the most important factor in their selection. Because of the small number of available teachers with this limited experience in Seattle, the group possessed some characteristics which were not common to the population for the major study in the fall. For example, the teachers were both secondary and elementary teachers, rather than being limited to the elementary level. The teachers were also of both sexes, and some of them had been trained in institutions outside of the State of Washington. None of the teachers in this pilot study received any of the experimental treatments designed for the major study.

The observers trained for this pilot study were a former secondary teacher of agriculture; a teacher trained for junior high classrooms, but who was a substitute teacher in the Seattle District in the intermediate grades; and a primary grade teacher who was a substitute teacher in the Seattle system. During the training period, the staff worked with other professional persons who they thought might be potential observers for the study, including supervisors of student teachers and consultants to other elementary teachers. In this limited experience, the staff decided that the substitute teachers recruited for this study were somewhat more trainable than the experienced supervisors. This belief was important to the judgements made later in attempting to recruit observers for the major study.

Training included seminar sessions with Dr. Harry L. Garrison, a Research Associate with the project, whose special field of competency was training of Seattle school personnel in the appraisal of classroom behavior. Dr. Garrison's previous research included the development of the Stanford Appraisal Guide. His major responsibility on this project was to train and supervise the teams of observers who visited the classrooms of the beginning teachers in the study. The process of training the observers included visits to classrooms of experienced, able Seattle teachers and to the classrooms of cadet teachers. Observers, in this way, were able to study a range of behavior for the different criteria used in the rating forms. A need that arose in the study was to develop materials for training observers which would provide for repeated study of exactly the same teacher behavior. To meet this need, the research staff developed video-taped recordings of teachers in the Pullman, Washington schools. These tapes were used repeatedly throughout the project as training materials for observers.

During this pilot study each of the 28 beginning teachers was visited by this team of three trained observers on three different occasions. The visits were four weeks apart and the two instruments under study were rotated equally among the observers and among the 28 classrooms.

A trial form of an attitude test was developed and tested with the 28 pilot study teachers. The attitude test was based upon the semantic differential concept of Osgood and others. All of the concepts, which in the staff's judgement were plausible objects of beginning teacher opinions or attitudes, were included in this trial instrument. Also, pairs of adjectives were included from the Osgood material which seemed applicable to these concepts in this context. This rather lengthy instrument was to be given each of the 28 subjects on each of the three occasions when the observers visited their classrooms. A copy of the trial form of the attitude scale, appears in Appendix A in underline.

The 28 teachers for the study met just before the first round of observations with the research staff, who explained to them the purposes and procedures of this pilot program. On that occasion each of the 28 teachers was given a questionnaire to fill out, furnishing the staff with information about each of the subject's personal and professional backgrounds. A copy of this questionnaire is shown in Appendix A. The purpose of this information, as was explained above, was to collect data which might be related to the classroom behavior of beginning teachers.

It was hoped by the staff that these analyses would indicate which two factors about the beginning teacher would be important as matching variables in the fall study.

A diary was kept of the experiences of the observers during training and during the actual observations. This anecdotal record was meant to serve as a guide to the staff in the later conduct of the actual experiment.

The data from the pilot study consisted then of judgements by three trained observers, based upon two different rating scales and made on three different occasions, as to the classroom behavior of 28 beginning Seattle teachers. The data also consisted of these 28 teachers' responses to the trial form of a Semantic Differential Test of attitudes towards teaching.

Results of the Pilot Study

Both of the instruments used to appraise classroom performance of the 28 teachers in the pilot study were judged administratively efficient and capable of high reliability. The coefficients of correlation between pairs of observers using the Ryans Scale and the Stanford Form appear in Tables 2 and 3.

The independent judgements of the three observers, when compared on the different criteria of each instrument, were gratifyingly highly related. It appeared to the research staff that the first two basic assumptions of the study were upheld by this pilot study experience.

The data concerning observer ratings on the two instruments did not indicate a clear preference for one of the two instruments. The three observers were unanimous in their opinions, however, that the Stanford appraisal guide was the simpler to administer and the more logical to

TABLE 2
OBSERVER CORRELATIONS[a] IN THE PILOT STUDY FOR THE STANFORD FORM

Observers	Round		
	1	2	3
1 & 2	.85	.84	.89
1 & 3	.75	.61	.95
2 & 3	.77	.79	.91

[a]Pearson product moment correlations.

TABLE 3
OBSERVER CORRELATIONS[a] IN THE PILOT STUDY FOR THE RYANS FORM

Observers	Round		
	1	2	3
1 & 2	.77	.72	.93
1 & 3	.58	.75	.77
2 & 3	.82	.90	.90

[a]Pearson product moment correlations.

use. Their feelings were that viewing the teaching act was apt to lead to a more objective judgement than studying the characteristics of the teacher, which was the focus of the Ryans scale.

On the basis of the observers' judgements and partly due to the fact that the Research Associate, Dr. Harry L. Garrison, had extensive experience with one of the scales, the Stanford appraisal guide was chosen as the basic instrument for classroom observations.

During the summer of 1965, Dr. Garrison, working with the observers from the pilot study, modified the Stanford scale somewhat to make it easier to administer and likely to be more consistently applied by different observers. The resulting modification of the Stanford guide is called the Seattle Teaching Performance Appraisal Guide. This became the basic instrument for assessing the behavior of the experimental subjects in the study. A copy of the Seattle Teaching Performance Appraisal Guide appears in Appendix A.

The experience of the staff in training the observers for the pilot study led to some new practices for preparing the observers who worked with the actual experiment. One unexpected finding was that observers were found to need repeated training throughout the study. The opinion of Dr. Garrison and the research staff was that the appraisal technique required the observer to develop some mental models of different kinds of classroom behavior of teachers. As the observer applied these models in appraising specific teaching acts, subtle changes in these mental models developed; and, with time, the criteria used by a particular observer became less and less the criteria with which he started the observation experience. Thus it seemed to the staff that repeated training, preferably with video-taped recordings, was necessary to maintain a consistent application of the instrument.

In the judgements of the observers and Dr. Garrison, the Stanford appraisal guide, as revised in the form of the Seattle appraisal guide, clearly was applicable to the elementary school classroom as well as to the secondary classroom. Because this form is based upon a model of a complete teaching performance, it was decided that it would be essential for observers see the teacher at the point where classes begin in the morning or after an intermission during the school day. The observers found, however, that all steps in the teaching act, as developed in the Stanford form, could be viewed in a typical classroom of an elementary teacher in approximately one hour.

The semantic differential trial form designed to test attitudes was administered three times to the 28 subjects in the pilot study. The responses of the subjects were tabulated for each concept and for each pair of adjectives designed to test attitudes regarding that concept. Table 25 in Appendix B shows the mean ratings of the 28 subjects and standard deviations of these means for each pair of adjectives, for each concept in the trial form.

It was the staff's judgement that the trial form of the attitude scale was too long. The beginning teachers found it tedious to complete and thus tended to mark the responses without sufficient care. In the interests of shortening and improving the scale the staff eliminated what seemed to be the least promising concepts and pairs of adjectives. Through inspection, the staff eliminated concepts on which responses tended to group near 4.0, and for which standard deviations were small. These kinds of patterns of responses suggested that attitude intensity was low. (4.0 is "neutral" on the seven point scale.) The staff also eliminated those items where means of responses were near 1.0 and where standard deviations were small.

These patterns suggested that there was little range of response, and variations among groups in the experiment would be small. Also, some adjective pairs consistently appeared to lack discrimination power, and these were eliminated. The staff also decided to use only concepts for which at least six adjective pairs appeared to be promising discriminators. One of the concepts did not meet these subjective criteria and was still kept in the final form of the scale. This concept was "teaching as a career." The staff left this one in the final form because it was most closely related to the principle concern of the project in this phase of the study. The revised form of the Semantic Differential Attitude Scale appears in Appendix A.

TABLE 4

PRODUCT MOMENT CORRELATIONS OF PILOT STUDY SUBJECTS' GRADE POINT WITH OBSERVER RATINGS OF THEIR TEACHING
(N=26)

Grade Point Averages	Correlation Coefficients With:	
	Ratings on Stanford Form	Ratings on Ryans Form
Cumulative Grade Point average in college	.06	-.68
GPA in Education Courses	.20	.19

TABLE 5

MEAN TOTAL SCORES ON TEACHING PERFORMANCE BY PILOT STUDY SUBJECTS ACCORDING TO COLLEGE ACADEMIC MAJOR

Major	N	Mean of Total Scores on Stanford Form	Mean of Total Scores on Ryans Form
Social Studies	12	107.75	212.25
Languages Arts	6	109.16	224.
Other	8	102.75	224.50

TABLE 6
TEACHING PERFORMANCES OF PILOT STUDY TEACHERS COMPARED TO GRADES IN STUDENT TEACHING

	At or Below Mean of Ratings on the Ryans Scale	Above Mean of Ratings on Ryans Scale
"A" in Student Teaching	7	7
"B" in Student Teaching	6	6

Chi Square = 0

	At or Below Mean Ratings on the Stanford Scale	Above the Mean Rating on the Stanford Scale
"A" in Student Teaching	6	8
"B" in Student Teaching	7	5

Chi Square for 2 x 2 fold table, corrected for continuity[a] = 1.39 (n.s)

---

[a] Henry E. Garrett, "Statistics in Psychology and Education", David McKay Co., Inc. 1962, p. 265.

TABLE 7
TEACHING PERFORMANCES OF PILOT STUDY SUBJECTS COMPARED TO GRADE LEVEL ASSIGNMENT

	Below Mean of Ryans Scale Ratings	Above Mean of Ryans Scale Ratings
Primary	7	7
Intermediate	3	5

Chi Square for 2 x 2 fold table, corrected continuity[a] = .87 (n.s.)

	Below Mean of Stanford Scale Ratings	Above Mean of Stanford Scale Ratings
Primary	7	7
Intermediate	4	4

Chi Square = 0

---

[a] Garrett, op. cit.

Table 26, Appendix B, shows a summary of the data on teacher behavior as related to general characteristics of the beginning teachers in the pilot study. No characteristics were found to be significantly associated with the level of teaching performance as determined by observer ratings. The design of the project called for a matching procedure in which clusters of four subjects would be matched according to two characteristics. As the data in Table 24 do not suggest two variables, the staff used their own judgements and those of members of the advisory committee to the project to make this selection. Grade in student teaching and primary or intermediate grade-levàl assignment were chosen because they were each two-factor variables, and would be simple to apply in matching. Also, the committee and staff judged them to be likely to be related to teaching performance in a larger sample than that of the pilot study.

Tables 4-7 show comparisons of pilot study subjects' characteristics with teaching performances.

One comparison, college grade point with Ryans scale ratings, appeared to be negatively significant. Because the other three comparisons with grade point approached zero correlation, the staff judged this significant correlation to be due to chance. Final ratings of teachers in the fall experiment were correlated with grade point, however, to check on this judgement.

The results of the pilot study then were:

1. The appraisals by trained observers of the classroom performances of teachers were judged to be administratively feasible and sufficiently reliable for purposes of this research.
2. The Stanford Appraisal Guide, as adapted, was judged easier to administer than the Ryans Scale, and more likely to be applied consistently by different observers.
3. The factors which were selected to be used in a matching procedure for selecting experimental subjects were grade level assignment and student teacher grade.

4. Thorough and repeated training of observers was judged essential for maintaining reliability in the use of the appraisal instruments.

5. The Semantic Differential Attitude Scale was revised and shortened according to the relative discriminating power of adjective pairs and concepts which were responded to by 28 beginning teachers.

Procedures in the Major Experiment

While the pilot study was being undertaken, the research staff met with school districts in the Seattle metropolitan area to arrange the conditions which were necessary for the experiment. Five of the largest school districts in the State of Washington agreed to participate in the study. These districts were the Seattle School District, Shoreline School District, Edmonds School District, Bellevue School District, and Lake Washington School District. To limit the factor of varying school district populations, schools in the central area of Seattle were not used for the study. Those school populations in Seattle which were used, and the four other districts, comprise outlying residential areas in metropolitan Seattle. It was not feasible to control the socio-economic backgrounds of the pupils in the study to a greater degree. The cooperation of these five districts was a major factor in the successful completion of this research. Indirect costs borne by these districts were a substantial contribution to the financing of the project.

During the summer of 1965 Dr. Harry Garrison worked with nine experienced teachers who were recruited to serve as observers for the study. The summer training period consisted of seminar discussions and practice in applying the Seattle appraisal guide to video-type recorded performances of classroom teachers.

During the same period of time the staff obtained the names of all beginning elementary teachers who had been hired by the five participating districts. Data on these teachers were also collected which included: the college from which the teacher graduated, the student teaching grade, the college grade point average, the school district's assignment of the new teacher, the college major study, and the marital status of this teacher. These names became the potential population from which the subjects of the study were selected.

Near the end of the summer of 1965, Dr. Garrison worked with supervisors from the five participating districts who were involved in the treatment of the Group I experimental subjects. These supervisors were to observe each beginning teacher in Group I for a period of approximately an hour and one-half twice during each week, using the Seattle appraisal guide, and then were to confer with the beginning teacher during that teacher's released time. The conference was to be based upon the supervisor's observation of the beginning teacher. The supervisors involved in this treatment were given similar training to that given the observers.

The research staff met with principals and other administrative staff members in each of the five participating school districts to explain the project and to indicate the nature of the support that was expected of the principal. Principals of teachers in Group II were asked to select experienced, able teachers which would be appropriate for the experimental subjects in Group II to visit during their released time.

In late August, the staff worked with the administrators of the five school districts to make it possible to juggle class enrollments so that Group III teachers would have the 25 per cent smaller class size which was required by the design of the study.

Just prior to the beginning of school the research staff selected the subjects from the population of new elementary teachers. The subjects selected met these criteria: They were women; they were June graduates of Washington teacher-education institution; this was their first teaching experience; and their assignment was limited to an elementary school with grades kindergarten through six. From this group, clusters of four teachers were matched on the bases of their respective grades in student teaching ("A" or "B"), and their grade level assignments (primary or intermediate). These clusters of four matched teachers were then randomly assigned to each of the four treatment groups.

The research staff then met with the selected beginning teachers in each of the five participating districts. At these meetings the staff explained the purposes of the study and the nature of each group's treatment. It was decided to be completely candid with the experimental subjects rather than to attempt to conceal the treatment of groups other than the ones in which they were participating. The advice of school administrators was that this could not be kept secret at any rate, and would probably be less threatening if the teachers knew the information they would find out anyway at the earliest possible time. The beginning teachers selected for the study were given the alternative of remaining in the study or dropping out. A few of the selected subjects asked to be dropped from the study. Last minute enrollment problems in the cooperating school districts made it impossible for some other selected teachers to participate. The final number of subjects at the beginning of the study in the fall was 105.

In each school district, substitute teachers were recruited to replace Group I and Group II teachers during one-quarter of each school day. Typically, the substitute teacher relieved the Group I and Group II teacher for a period of time, either after lunch until the afternoon break in the school day, or from that point until the end of the school day. The same substitute teacher remained with each Group I or II teacher throughout the period of the experiment. Substitute teachers did not take up their duties, however, until after the Group I and II teachers had been visited the first time by teams of trained observers.

Group III teachers were assigned to classrooms which the administration had been able to limit to the smaller enrollment required under the design of the study. Group IV teachers did not receive any of the treatments planned for this experiment, but did participate in the regular orientation procedures held by the five districts for all their beginning teachers.

A final period of training was given the nine observers who would visit the beginning subjects. This final training included visits to the classrooms of experienced Seattle teachers, and to classrooms where cadet teachers were doing their practice teaching. These observers also reviewed video tapes of experienced teachers. The first round of observations began one week after the fall term opened in the five districts and continued for a two and one-half week period. During this round of observations, as in subsequent rounds, each classroom teacher in the four groups was visited by a team of three observers. The observers made their judgements independent of each other, and recorded those judgements on the Seattle appraisal guide while they were actually in the classroom.

Prior to this visit each teacher had received a letter informing them of the impending visit, and also asking them to fill out an enclosed copy of the Semantic Differnntial Attitude Scale. Upon leaving the classroom teacher after the observation, the observers picked up the completed Semantic Differential Attitude Scale.

Between the first and second round of observations the nine observers went through a brief period of retraining, and were reassigned to different teachers and as members of different teams of observers. The plan of the project was to rotate observers so that there would be a minimum number of times when any given observer worked with the same observer or saw the same beginning teacher. (See Table I)

After this first round of observations, experimental treatments for Groups I and II subjects were begun. Group III teachers had received the experimental treatment from the opening of school. This was necessary because there was no administrative method for reassigning pupils after school started to meet the experimental conditions of a smaller class size as required in the study. Thus a limitation in the study was that the Group III teachers received from two to four weeks more treatment than did the Group I and II teachers.

Observers visited the beginning teachers at one of three times:

at the beginning of school in the morning, after classes reconvened following the morning recess, or at the beginning of classes after lunch. The team remained with the classroom teacher until they had observed, to each observer's satisfaction, the entire teaching act which was the model for the Seattle appraisal form. If a particular teaching behavior could not be observed by an observer, he was instructed to mark that behavior zero.

Four and one-half weeks after the first round of observations, the teams made a second round and also administered the Semantic Differential Attitude Test a second time. This procedure was repeated, always with additional training of the observers, a third and a fourth time. Final observations were completed in January, just before the end of the fall school term. The schedule of observations was made so that the interval between observations was the same for each beginning teacher in the study. During the fall term, because of illness or other reasons, the experimental group was reduced to 96. This rather large loss of subjects was due partly to the fact that subjects were matched in clusters of four. When one subject was lost, data for all four matched subjects in the cluster had to be dropped from the final analysis.

The data collected from the experimental phase of the project included observations by three different teams of trained observers of beginning teachers who were given four different treatments. These observation data were ratings of ten different teaching behaviors appraised by the three observers on four different administrations. The data also included four different administrations of an attitude scale to these beginning teachers. These data were then analyzed by computers in the Research Office of the State Superintendent of Public Instruction.

CHAPTER V

ANALYSIS AND DISCUSSION OF THE DATA

Introduction

The main effect to be studied was the experimental treatments of three groups of matched beginning teachers compared to a fourth matched group on the basis of their classroom performances. The data collected for this analysis were the judgements of observer teams as to ten different teaching behaviors of these subjects. These judgements were made four times at approximately one-month intervals.

The first step in the analysis of the main effect of the study was an assessment of the reliability of the observers' ratings. The independent judgements of observers who rated the same teacher were compared as a test of reliability. To see if the ten different teaching behaviors rated were in fact different, the combined ratings of observers for each behavior were correlated with each of the other behaviors.

Then the observer ratings of the four treatment groups were analyzed to see if there were differences in teaching performances among the groups. These analyses of group differences were made for each of the four rounds of observer ratings.

A secondary effect of the study was to be the apparent effects of treatments upon the attitudes of the beginning teachers. Teachers in the four treatment groups responded to an attitude scale based upon the semantic differential technique as reported by Osgood and revised on the basis of pilot study results. The attitude scale was administered four times at monthly intervals. The responses of the four groups were studied to determine if there were differences among the four groups in attitudes towards different concepts associated with teaching.

The total group's ratings on the four observation rounds were compared with grade point averages in college and, separately, with grade point in Education courses.

Analysis of the Observation Rating Procedures

Teams of three observers visited teachers in each of the four treatment groups. The observers made their judgements of ten different teaching behaviors independently. There were a total of nine observers. In the four rounds of observations, observer assignments were rotated so that a different team judged each teacher on each of the four rounds. Table 8 shows correlation coefficients of the three observers' ratings in each team for each of the four rounds. In order to compare three observer scores, Pearson product moment coefficients, which were obtained for each pair of observers, were converted to Fisher Z scores.[1] Product moment correlations between pairs of observers are reported in Appendix B.

As reported in Table 8, correlations of observer judgements ranged from .29 to .88. The total correlation of all observers on each round varied from .38 to .82. The degree to which observers judged teachers the same way was less than reported in the pilot study. This lesser amount of agreement may have been due to the greater number of observers (nine compared to three), or to the changes in the evaluation instrument. The instrument

---

[1] Henry F. Garrett, "Statistics in Psychology and Education", New York: David McKay Co., Inc., 1962, pages 134, 172.

TABLE 8

OBSERVER CORRELATIONS[a] BY TEAM BY ROUND BY BEHAVIOR

Observer Team	Behaviors										
	1	2	3	4	5	6	7	8	9	10	Overall
					Round 1						
1	.40	.45	.52	.52	.35	.52	.49	.54	.49	.29	.46
2	.52	.55	.66	.63	.61	.65	.65	.56	.39	.46	.57
3	.63	.67	.58	.63	.54	.59	.61	.55	.65	.57	.59
					Round 2						
1	.56	.75	.68	.59	.74	.67	.66	.62	.72	.72	.67
2	.69	.74	.75	.75	.62	.76	.78	.71	.61	.56	.72
3	.24	.24	.24	.20	.32	.31	.29	.29	.31	.23	.27
					Round 3						
1	.80	.87	.78	.81	.84	.88	.80	.81	.85	.81	.82
2	.28	.28	.36	.35	.53	.52	.53	.29	.29	.33	.38
3	.64	.60	.54	.58	.61	.58	.67	.60	.54	.58	.59
					Round 4						
1	.29	.49	.45	.38	.63	.61	.49	.40	.40	.47	.47
2	.76	.79	.77	.77	.75	.77	.76	.75	.80	.78	.77
3	.81	.79	.86	.81	.86	.86	.84	.81	.82	.80	.82

[a] Correlations are from converted Fisher Z scores of Pearson r's

used in the experiment was adapted from the Stanford scale used in the pilot study. A third possible source of variation might be the interaction of observers and the instrument. Inspection of the data on observer agreement revealed that two of the nine observers accounted for most of the variations in observers' judgements. When the ratings in which the two observers participated are omitted (Table 9) the remaining coefficients of correlation, especially on the fourth round, are similar to those obtained in the pilot study.

In spite of the relatively low degree of observer agreement, calculations based upon the Spearman-Brown reliability formula[1] produced coefficients sufficiently high in the judgement of the project staff to warrant further treatment of the observer data. (See Table 10)

Inter-item correlations were computed for pairs of the ten behaviors. These correlations were calculated for each of the four rounds of observations. The coefficients of correlation between teaching performances based upon mean ratings by observers in Round 4 are reported in Table 11. The very high correlations obtained suggest that either different behaviors are closely related in actual practice or observers tended to be influenced by some behaviors when rating others.

The evaluation instrument was conceived of as ten behaviors constituting four steps in a teaching cycle. These steps were (a) combinations of behaviors 1 and 2, (b) behaviors 3, 4 and 5, (c) behaviors 6, 7 and 8, and (d) behaviors 9 and 10. Inspection of the correlations between behaviors suggests, however, that there is little justification for seeing these combinations as distinct steps in the teaching process, and considerable

[1] Garrett, page 343.

TABLE 9

CORRELATIONS OBTAINED WHEN TWO OBSERVERS' RATINGS WERE OMITTED

Behavior	Round 1	Round 2	Round 3	Round 4
1	.61	.63	.73	.79
2	.67	.75	.76	.80
3	.59	.71	.68	.81
4	.61	.68	.71	.80
5	.56	.69	.76	.81
6	.59	.72	.77	.82
7	.64	.72	.74	.80
8	.62	.67	.72	.68
9	.59	.67	.73	.81
10	.53	.65	.71	.79
Overall	.60	.69	.72	.80

TABLE 10

RELIABILITY OF OBSERVER APPRAISALS BY TEAM BY ROUND BY BEHAVIOR

Observer Team	Behavior										
	1	2	3	4	5	6	7	8	9	10	Overall
						Round 1					
1	.67	.71	.77	.77	.62	.77	.74	.78	.74	.55	.72
2	.77	.79	.85	.84	.82	.85	.85	.79	.66	.72	.80
3	.84	.86	.81	.84	.78	.81	.82	.79	.85	.80	.81
						Round 2					
1	.79	.90	.86	.81	.90	.86	.85	.84	.88	.88	.86
2	.87	.90	.90	.90	.84	.87	.85	.88	.82	.79	.88
3	.49	.49	.49	.43	.59	.57	.55	.55	.57	.47	.53
						Round 3					
1	.92	.95	.92	.93	.94	.96	.92	.93	.95	.93	.93
2	.54	.54	.63	.62	.77	.77	.77	.55	.55	.60	.64
3	.84	.82	.78	.81	.82	.81	.86	.82	.78	.81	.81
						Round 4					
1	.55	.74	.71	.65	.84	.82	.74	.67	.67	.73	.73
2	.87	.92	.91	.91	.90	.91	.91	.90	.92	.92	.91
3	.93	.92	.94	.93	.94	.94	.94	.93	.93	.92	.93

TABLE 11

INTER-ITEM CORRELATIONS OF ROUND 4 RATINGS ON EACH OF THE TEN PERFORMANCES, INDIVIDUALLY AND IN COMBINATIONS

Teaching Performances	TEACHING PERFORMANCES														
	1	2	3	4	5	6	7	8	9	10	1 & 2	3,4 5	6,7 8	9 & 10	1-10
1. Suitability of Goals	1.00	.88	.91	.95	.91	.91	.89	.89	.91	.91	.97	.95	.92	.92	.96
2. Acceptance of Goals		1.00	.87	.88	.89	.89	.88	.89	.91	.88	.97	.91	.91	.91	.94
3. Use of Resources			1.00	.92	.90	.88	.88	.92	.88	.89	.92	.96	.92	.90	.95
4. Choosing the Plan				1.00	.94	.92	.92	.92	.91	.90	.94	.98	.94	.92	.97
5. Organizing Class					1.00	.96	.94	.92	.91	.90	.92	.97	.96	.92	.97
6. Classroom Control						1.00	.95	.91	.91	.90	.93	.94	.98	.92	.97
7. Class Climate							1.00	.91	.89	.89	.91	.94	.98	.91	.96
8. Active Participation								1.00	.89	.88	.92	.95	.97	.90	.96
9. Measure Results									1.00	.92	.93	.92	.92	.98	.95
10. Using Measures										1.00	.92	.92	.91	.98	.95
11. 1 & 2											1.00	.95	.94	.95	.98
12. 3, 4, 5												1.00	.97	.94	.59
13. 6, 7, 3													1.00	.93	.99
14. 9 & 10														1.00	.97
15. All (1-10)															1.00

justification for interpreting all the behaviors as closely related aspects of a general teaching performance.

Observers were instructed to score zero for a teaching behavior which they felt they were unable to observe. One test of the usefulness of the instruments was believed to be the percentage of zero scores. Only 18 of 1260 behaviors, 1.42 per cent, were scored zero. These zero ratings were judged to be an insignificant limitation to the usefulness of the instrument.

Finally, each of the nine observers and two members of the research staff, who were occasionally used as substitute observers (one of whom was an observer in the pilot study), were questionnaired as to their opinions about various aspects of the instrument. A summary of their responses to the questions which were asked appears in Table 12. These tabulations show that the observers were nearly unanimous in their belief that the instrument was a fair and valid basis for evaluating teaching, and that each teaching performance represented a distinct behavior which they felt they had no difficulty in isolating during their appraisals of the beginning teachers.

Analysis of the observers' ratings and a questionnaire to the observers resulted in the following judgements by the project staff concerning the observation procedure:

1. Observer agreement was less than anticipated, but was judged to be adequately reliable for purposes of further analysis of the data.
2. The relatively low agreement among observers probably produced considerable variability of the ratings within treatment groups.

TABLE 12
OBSERVER EVALUATION OF TEACHING BEHAVIORS

Question	Behaviors									
	1	2	3	4	5	6	7	8	9	10
Can an experienced teacher demonstrate this behavior pattern?.	4.7	4.7	4.9	4.9	5.0	4.9	5.0	4.7	4.7	4.6
Can an experienced teacher observe the behavior pattern when demonstrated	4.4	4.9	3.5	4.6	5.0	4.9	4.9	4.9	4.4	4.2
Does the pattern have psychological validity?.	4.9	4.9	5.0	4.6	4.9	4.9	5.0	4.9	4.9	4.9
Does the pattern have sociological validity? .	4.9	4.8	4.9	4.9	4.9	4.9	4.9	4.9	4.7	4.7
Is this behavior pattern really necessary for successful teaching?	4.9	4.9	4.9	5.0	4.9	4.9	5.0	5.0	4.9	4.9
Can differences in performance level of this behavior pattern be recognized by an experienced teacher?.	4.4	3.9	3.9	4.5	4.9	4.9	4.9	4.4	4.2	3.6
Does this behavior pattern permit the teacher to develop his own unique style?	5.0	4.9	4.9	5.0	5.0	4.9	5.0	4.6	4.7	4.5

a Range of each item was 1-5; 1 = "rarely or doubtful"; 5 = "very much so".

3. High correlations between mean scores of different teaching behaviors suggested that the different scores should be treated as different approaches to rating one general complex of teaching performances.

4. The observers supported the evaluation instrument and the procedures for appraising teachers. This was judged to be supporting evidence that the procedures were valid.

Analysis of Classroom Performance

The observer teams made their first visits to the teachers in the four different treatment groups beginning with the second week of school and continuing through the fourth week. The mean ratings of these observer teams for each of the ten teaching behaviors are reported for each group in Table 13. The same table also reports the F ratios in an analysis of variance. No differences which were significant were found among the groups for any of the ten behaviors. Except for behavior number 1, the F ratios, in fact, approached zero. Apparently the matching procedures and selection techniques were effective in equating the four groups at the outset of the study. Inspection of the mean scores shows that the average levels of performance of the subjects were somewhat below the average to be expected of a theoretical teaching population universe, i.e. 4.0 on a seven point scale.

The design of the experiment called for a procedure of adjusting such differences as would appear in this initial round of observation through an analysis of covariance. In the covariance analysis the means on the initial round of observations were to be used as the control variable in the subsequent analysis of differences among groups on these later observer ratings. This analysis of covariance followed six steps suggested by Lindquist.[1]

---

[1] E. F. Lindquist "Design and Analysis of Experiments in Psychology and Education", Houghton Mifflin Co., Boston 1956, pages 332-333.

TABLE 13

MEAN PERFORMANCE RATINGS OF THE FOUR TREATMENT GROUPS ON FIRST ROUND OF OBSERVATION

Teaching Performances	Treatment Group Means				F Ratios
	Group 1	Group 2	Group 3	Group 4	
1. Suitability of Goals	2.76	3.13	3.00	2.54	.78
2. Acceptance of Goals	2.65	2.63	2.80	2.61	.07
3. Explanation of Resources	2.65	2.90	3.04	2.78	.23
4. Selecting the Plan	2.57	2.76	2.81	2.59	.13
5. Organizing the Class	2.74	2.72	2.82	2.96	.12
6. Classroom Control	2.72	2.89	2.89	2.85	.07
7. Classroom Climate	2.61	2.70	2.87	2.80	.12
8. Active Participation	3.04	2.89	3.09	2.93	.09
9. Measuring Achievement	2.72	2.92	2.81	2.48	.04
10. Using Measurement	2.70	2.82	2.70	2.47	.23
11. Performances 1 and 2	2.70	2.88	2.90	2.57	. 6
12. Performances 3, 4 and 5	2.65	2.79	2.89	2.78	.09
13. Performances 6, 7 and 8	2.79	2.83	2.95	2.86	.05
14. Performances 9 and 10	2.72	2.87	2.76	2.48	.30
15. Performances 1 - 10	2.72	2.84	2.88	2.70	.09

Observers were re-assigned to different teams and went through a period of retraining before making a second round of visits to the beginning teachers. This second round took place four and a half weeks after the first round. The ratings of observers for each group and for each teaching behavior are reported in Table 14. Observers judged the entire population consisting of all four groups to be inferior in performances when compared to the initial round of observations. On the 40 mean ratings, 38 are lower than the mean ratings made in the first round. Inspection of these data suggests that there are no differences among treatment groups in the patterns of change.

The analysis of covariance is shown in Table 15. As expected from the inspection of data, no significant differences were found among the four groups for any of the ten behaviors.

A third round of observations was made following the same procedures which were carried out in the second round. Observer teams' judgements are reported for each of the ten behaviors in Table 16. Inspection of this table shows a different tendency for the mean scores to vary than was found in the second round. Groups I and II show consistent improvement, while Groups III and IV remain at relatively low levels of teaching performance. All of the mean ratings for Groups I and II show gains over their observer ratings on the second round. In the case of Group I, nine of the ten behaviors are rated higher than they were rated on the first round. In the case of Group II, seven of the ten behaviors were superior to round one. This consistent pattern of changes was not considered to be explainable in terms of pure chance.

The analysis of covariance is shown in Table 17. No significant differences were found among the four groups for any of the ten behaviors.

TABLE 14

MEAN SCORES OF FOUR TREATMENT GROUPS ON OBSERVATION ROUNDS 1 AND 2

	Teaching Abilities	GROUP I			GROUP II			GROUP III			GROUP IV		
		Rounds		Diff.	Rounds		Diff.	Rounds		Diff.	Rounds		Diff.
		1	2		1	2		1	2		1	2	
1.	Suitability of Goals	2.76	2.50	-.26	3.13	2.65	-.48	3.00	2.65	-.35	2.54	2.31	-.23
2.	Student Acceptance of Goals	2.65	2.14	-.51	2.63	2.45	-.18	2.80	2.72	-.08	2.61	2.20	-.41
3.	Use of Resources	2.65	2.33	-.32	2.90	2.46	-.44	3.04	2.59	-.45	2.78	2.04	-.74
4.	Selecting the Plan	2.57	2.44	-.13	2.76	2.59	-.17	2.81	2.44	-.37	2.59	2.22	-.37
5.	Organizing Class	2.74	2.52	-.22	2.72	2.78	+.06	2.82	2.67	-.15	2.96	2.48	-.48
6.	Class Control	2.72	2.39	-.33	2.89	2.60	-.29	2.89	2.68	-.21	2.85	2.37	-.48
7.	Classroom Climate	2.61	2.28	-.33	2.70	2.63	-.07	2.87	2.65	-.22	2.80	2.43	-.37
8.	Student Participation	3.04	2.59	-.45	2.89	2.74	-.15	3.09	2.81	-.28	2.93	2.46	-.47
9.	Measuring Achievement	2.72	2.10	-.62	2.92	2.47	-.45	2.81	2.43	-.38	2.48	2.21	-.27
10.	Using Measurement	2.70	1.85	-.85	2.82	2.45	-.37	2.70	2.20	-.50	2.47	2.14	-.33
11.	Performances 1 and 2	2.70	2.32	-.38	2.88	2.55	-.33	2.90	2.69	-.21	2.57	2.26	-.31
12.	Performances 3, 4 & 5	2.65	2.43	-.22	2.79	2.61	-.18	2.89	2.57	-.32	2.78	2.25	-.53
13.	Performances 6, 7 & 8	2.79	2.42	-.37	2.83	2.66	-.17	2.95	2.72	-.23	2.86	2.42	-.44
14.	Performances 9 & 10	2.72	1.98	-.74	2.87	2.46	-.41	2.76	2.32	-.44	2.48	2.18	-.30
15.	Performances 1 - 10	2.72	2.32	-.40	2.84	2.58	-.26	2.88	2.59	-.29	2.70	2.29	-.41

TABLE 15

ANALYSIS OF COVARIANCE: SIGNIFICANCE OF DIFFERENCES OF MEAN SCORES AMONG THE FOUR GROUPS, EACH OF 18 SUBJECTS, OBTAINED ON THE SECOND ROUND OF OBSERVATIONS (ROUND 1 OBSERVATION RATINGS USED AS COVARIABLE)

Source	Sums of Squares for Round 2	Sums of Cross-Products	Sums of Squares for Round 1	df
1. Teaching Performance: Suitability of Goals				
Treatments (T)	1.35	2.17	3.72	3
Subjects (S)	22.57	17.10	45.90	17
T x S (Error: E)	95.68	21.54	81.51	51
	F ratio = .083 (n.s.)			
2. Teaching Performance: Student Acceptance of Goals				
Treatments (T)	3.79	- .01	.38	3
Subjects (S)	38.37	1.61	56.49	17
T x S (Error: E)	93.98	- 2.10	95.42	51
	F ratio = .670 (n.s.)			
3. Teaching Performance: Use of Resources				
Treatments (T)	3.07	.40	1.51	3
Subjects (S)	29.85	1.43	49.19	17
T x S (Error: E)	103.77	.70	109.52	51
	F ratio = .493 (n.s.)			
4. Teaching Performance: Selecting the Plan				
Treatments (T)	1.27	.57	.78	3
Subjects (S)	33.86	16.13	49.78	17
T x S (Error: E)	97.13	23.94	99.64	51
	F ratio = .191 (n.s.)			

TABLE 15--Continued

Source	Sums of Squares for Round 2	Sums of Cross-Products	Sums of Squares for Round 1	df
5. Teaching Performance: Organizing the Class				
Treatments (T)	1.02	- .49	.65	3
Subjects (S)	42.49	32.26	68.06	17
T x S (Error: E)	111.44	19.80	91.07	51
	F ratio = .196 (n.s.)			
6. Teaching Performance: Class Control				
Treatments (T)	1.31	.46	.33	3
Subjects (S)	49.24	34.35	68.60	17
T x S (Error: E)	105.44	22.84	81.27	51
	F ratio = .182 (n.s.)			
7. Teaching Performance: Classroom Climate				
Treatments (T)	1.67	.69	.70	3
Subjects (S)	57.74	29.04	57.99	17
T x S (Error: E)	109.33	30.03	100.33	51
	F ratio = .219 (n.s.)			
8. Teaching Performance: Student Participation				
Treatments (T)	1.31	.31	.49	3
Subjects (S)	39.07	19.27	51.10	17
T x S (Error: E)	104.48	17.72	94.08	51
	F ratio = .198 (n.s.)			

TABLE 15--Continued

Source	Sums of Squares for Round 2	Sums of Cross-Products	Sums of Squares for Round 1	df
9. Teaching Performance: Measuring Achievement				
Treatments (T)	1.67	1.19	1.87	3
Subjects (S)	25.77	14.37	46.12	17
T x S (Error: E)	94.90	15.91	82.66	51
F ratio = .228 (n.s.)				
10. Teaching Performance: Using Measurement				
Treatments (T)	3.22	.71	1.16	3
Subjects (S)	29.92	16.64	43.74	17
T x S (Error: E)	86.27	8.61	86.57	51
F ratio = .602 (n.s.)				
11. Step One: Performances 1 and 2				
Treatments (T)	8.52	6.52	5.10	3
Subjects (S)	113.11	77.78	196.93	17
T x S (Errors: E)	355.02	80.39	335.95	51
F ratio = .286 (n.s.)				
12. Step Two: Performances 3, 4 and 5				
Treatments (T)	13.05	2.90	4.51	3
Subjects (S)	300.19	184.52	484.23	17
T x S (Error: E)	883.72	197.49	848.68	51
F ratio = .237 (n.s.)				

TABLE 15--Continued

Source	Sums of Squares for Round 2	Sums of Cross-Products	Sums of Squares for Round 1	df
13. Step Three: Performances 6, 7 and 8				
Treatments (T)	11.73	3.30	2.29	3
Subjects (S)	413.85	244.81	509.53	17
T x S (Error: E)	914.68	213.95	758.12	51
	F ratio = .196 (n.s.)			
14. Step Four: Performances 9 and 10				
Treatments (T)	9.11	3.81	5.92	3
Subjects (S)	107.12	61.54	178.13	17
T x S (Error: E)	350.03	48.49	331.50	51
	F ratio = .394 (n.s.)			
15. All Performances (total)				
Treatments (T)	144.11	76.91	42.92	3
Subjects (S)	3287.88	2095.80	5023.95	17
T x S (Error: E)	9042.52	2047.57	8152.93	51
	F ratio = .210 (n.s.)			

TABLE 16

MEAN SCORES OF FOUR TREATMENT GROUPS ON OBSERVATION ROUNDS 1 AND 3

		GROUP I			GROUP II			GROUP III			GROUP IV		
		Rounds		Diff.	Rounds		Diff.	Rounds		Diff.	Rounds		Diff.
	Teaching Abilities	1	3		1	3		1	3		1	3	
1.	Suitability of Goals	2.76	3.06	+.30	3.13	3.11	-.02	3.00	2.26	-.74	2.54	2.51	-.03
2.	Student Acceptance of Goals	2.65	2.93	.28	2.63	2.80	.17	2.80	2.20	-.60	2.61	2.24	-.37
3.	Use of Resources	2.65	2.96	.31	2.90	2.74	.26	3.04	2.26	-.78	2.78	2.45	-.33
4.	Selecting the Plan	2.57	2.91	.34	2.76	2.85	.09	2.81	2.33	-.48	2.59	2.41	-.18
5.	Organizing the Class	2.74	2.96	.22	2.72	3.19	.47	2.82	2.44	-.38	2.96	2.43	-.53
6.	Class Control	2.72	2.94	.22	2.89	2.94	.05	2.89	2.35	-.54	2.85	2.41	-.44
7.	Classroom Climate	2.61	2.78	.17	2.70	3.11	.41	2.87	2.33	-.54	2.80	2.33	-.47
8.	Student Participation	3.04	3.17	.13	2.89	3.08	.19	3.09	2.56	-.53	2.93	2.47	-.46
9.	Measuring Achievement	2.72	2.80	.08	2.92	2.93	.01	2.81	2.33	-.48	2.48	2.22	-.26
10.	Using Measurement	2.70	2.65	-.05	2.82	2.82	.00	2.70	2.03	-.67	2.47	1.93	-.54
11.	Performances 1 and 2	2.70	2.99	.29	2.88	2.96	.08	2.90	2.23	-.67	2.57	2.38	-.19
12.	Performances 3, 4 & 5	2.65	2.94	.29	2.79	2.93	.14	2.89	2.35	-.54	2.78	2.43	-.35
13.	Performances 6, 7 & 8	2.79	2.96	.17	2.83	3.04	.21	2.95	2.41	-.54	2.86	2.40	-.46
14.	Performances 9 & 10	2.72	2.73	.01	2.87	2.87	.00	2.76	2.18	-.58	2.48	2.08	-.40
15.	Performances 1 - 10	2.72	2.92	.20	2.84	2.96	.12	2.88	2.31	-.56	2.70	2.34	-.36

TABLE 17

ANALYSIS OF COVARIANCE: SIGNIFICANCE OF DIFFERENCES OF MEAN SCORES AMONG THE FOUR GROUPS, EACH OF 18 SUBJECTS, OBTAINED ON THE THIRD ROUND OF OBSERVATIONS (ROUND 1 OBSERVATION RATINGS USED AS COVARIABLE)

Source	Sums of Squares for Round 3	Sums of Cross Products	Sums of Squares for Round 1	df
1. Teaching Performance: Suitability of Goals				
Treatments (T)	9.37	1.34	3.72	3
Subjects (S)	29.85	17.64	45.90	17
T x S (Error: E)	86.69	17.75	81.51	51
F ratio = 1.80 (n.s.)				
2. Teaching Performance: Student Acceptance of Goals				
Treatments (T)	7.53	.16	.38	3
Subjects (S)	37.58	2.24	56.49	17
T x S (Error: E)	110.04	- 3.09	95.42	51
F ratio = 1.141 (n.s.)				
3. Teaching Performance: Use of Resources				
Treatments (T)	5.19	- .32	1.51	3
Subjects (S)	37.00	- 1.10	49.19	17
T x S (Error: E)	88.17	- .59	109.52	51
F ratio = .983 (n.s.)				
4. Teaching Performance: Selecting the Plan				
Treatments (T)	4.73	- .57	.78	3
Subjects (S)	32.18	10.14	49.78	17
T x S (Error: E)	97.18	17.27	99.64	51
F ratio = .878 (n.s.)				

TABLE 17--Continued

Source	Sums of Squares for Round 3	Sums of Cross-Products	Sums of Squares for Round 1	df
5. Teaching Performance: Organizing the Class				
Treatments (T)	7.84	- 1.87	.65	3
Subjects (S)	38.29	10.99	68.06	17
T x S (Error: E)	115.64	14.67	91.07	51
F ratio = 1.233 (n.s.)				
6. Teaching Performance: Class Control				
Treatments (T)	5.77	- .68	.33	3
Subjects (S)	39.16	12.64	68.60	17
T x S (Error: E)	111.60	18.97	81.27	51
F ratio = .949 (n.s.)				
7. Teaching Performance: Classroom Climate				
Treatments (T)	7.71	- 1.71	.70	3
Subjects (S)	40.40	15.49	57.99	17
T x S (Error: E)	104.45	25.60	100.33	51
F ratio = 1.459 (n.s.)				
8. Teaching Performance: Student Participation				
Teeatments (T)	6.77	- .25	.49	3
Subjects (S)	37.57	20.09	51.10	17
T x S (Error: E)	104.14	22.53	94.08	51
F ratio = 1.173 (n.s.)				

TABLE 17--Continued

Source	Sums of Squares for Round 3	Sums of Cross-Products	Sums of Squares for Round 1	df
9. Teaching Performance: Measuring Achievement				
Treatments (T)	6.37	2.36	1.87	3
Subjects (S)	26.77	15.61	46.12	17
T x S (Effort: E)	87.88	15.92	82.66	51
F ratio = 1.076 (n.s.)				
10. Teaching Performance: Using Measurement				
Treatments (T)	10.59	2.77	1.16	3
Subjects (S)	42.31	22.23	43.74	17
T x S (Error: E)	92.54	14.67	86.57	51
F ratio = 1.778 (n.s.)				
11. Step One: Performances 1 and 2				
Treatments (T)	32.99	.72	5.10	3
Subjects (S)	129.31	55.29	196.93	17
T x S (Error: E)	371.53	83.16	335.95	51
F ratio = 1.564 (n.s.)				
12. Step Two: Performances 3, 4 and 5				
Treatments (T)	49.13	- 10.65	4.51	3
Subjects (S)	303.12	100.26	484.23	17
T x S (Error: E)	814.23	147.53	848.68	51
F ratio = 1.222 (n.s.)				

TABLE 17-Continued

Source	Sums of Squares for Round 3	Sums of Cross-Products	Sums of Squares for Round 1	df
13. Step Three: Performances 6, 7 and 8				
Treatments (T)	57.71	- 9.05	2.29	3
Subjects (S)	334.08	139.35	509.53	17
T x S (Error: E)	882.22	190.38	758.12	51
F ratio = 1.304 (n.s.)				
14. Step Four: Performances 9 and 10				
Treatments (T)	33.38	10.27	5.92	3
Subjects (S)	130.84	75.89	178.13	17
T x S (Error: E)	339.15	62.06	331.50	51
F ratio = 1.501 (n.s.)				
15. All Performances (total)				
Treatments (T)	671.79	- 18.37	42.92	3
Subjects (S)	3528.96	1463.23	5023.95	17
T x S (Error: E)	8453.71	1875.44	8152.93	51
F ratio = 1.418 (n.s.)				

Apparently the total variance among the four groups was not sufficiently greater than the variance within each of the four groups. Contributing to this within-group variance was the relatively low amount of agreement among observers which has been discussed. Also the means for the total of all behaviors for each group ranged from 2.34 to 2.96 on a 7-point scale. The ratings were thus depressed within the 7-point scale to a relatively small range, limiting the potential amount of variance for this analysis.

The fourth round of observations took place in January, the last month of the fall term. The means for each of the four groups on the ten behaviors compared to the respective means of these groups on the first round of observations is reported in Table 18. The three experimental groups (I, II, and III) on this series of observations were rated as superior to Group IV in 28 of 30 comparisons. The degree of superiority on the total teaching performance of Group I over Group IV was 31 per cent, for Group II over Group IV was 26 per cent, and for Group III over Group IV was 58 per cent. These percentages are based upon the net gains of groups compared to the Group IV mean rating. In these ratings the lowest possible score was 1.0 To calculate the true relationship of the net gains of the groups to the mean rating of Group IV on all ten behaviors the actual mean rating of Group IV, 2.55, was converted to 1.55. The graph in Figure 1 shows the actual relationship of these scores.

Group III was superior to all groups and considered superior to Groups I and II by about the same degree these two groups were considered the superior of Group IV. The changes in rated performances on the part of Group III between round 3 and round 4 were the greatest degrees of change noted by observers in the entire study.

In terms of net change from the beginning of the study to the end

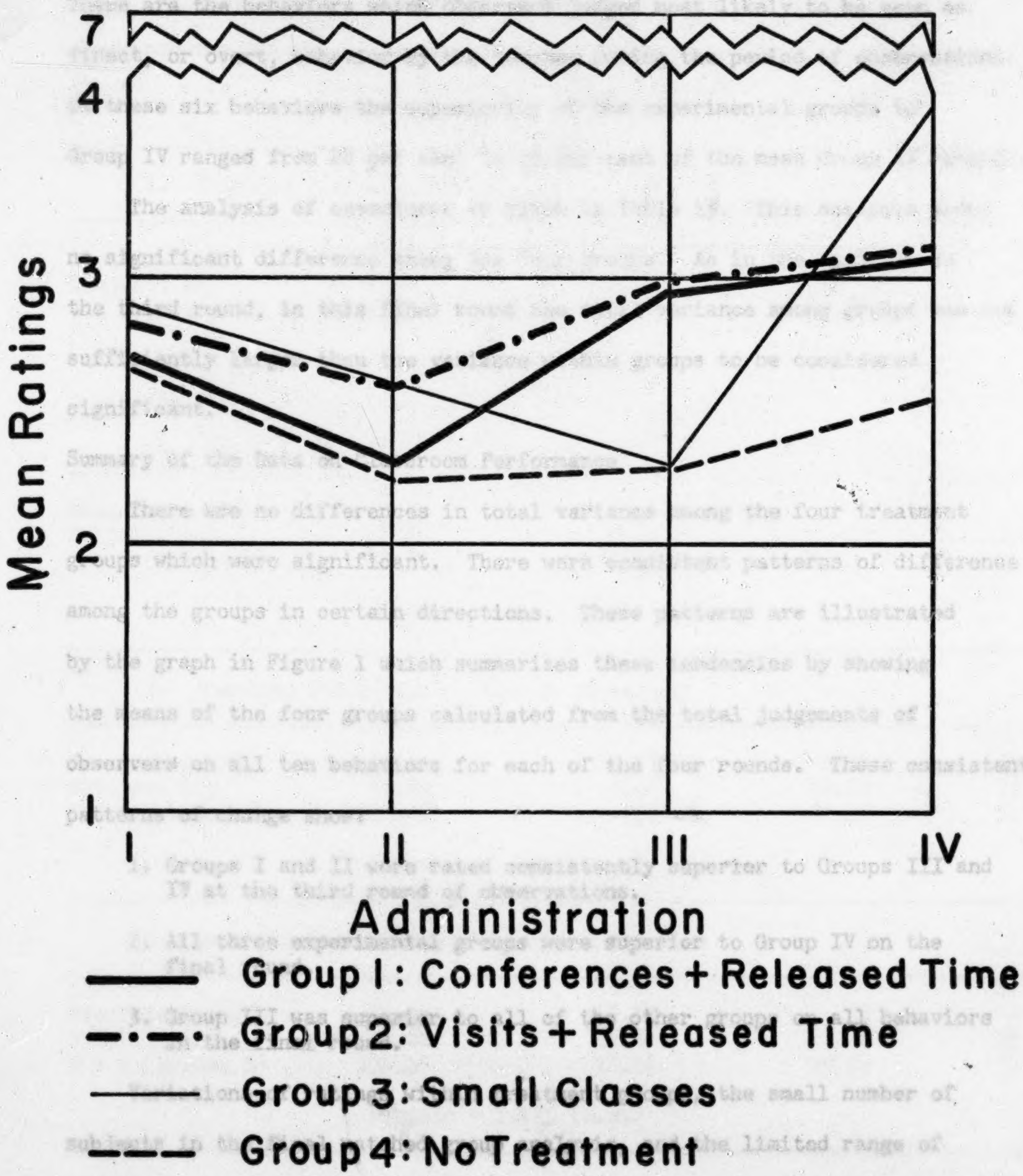
TABLE 18

MEAN SCORES OF FOUR TREATMENT GROUPS ON OBSERVATION ROUNDS 1 AND 4

		GROUP I			GROUP II			GROUP III			GROUP IV		
		Rounds		Diff.	Rounds		Diff.	Rounds		Diff.	Rounds		Diff.
	Teaching Abilities	1	4		1	4		1	4		1	4	
1.	Suitability of Goals	2.76	3.22	.46	3.13	3.20	.07	3.00	3.54	.54	2.54	2.61	.07
2.	Student Acceptance of Goals	2.65	3.06	.41	2.63	3.04	.41	2.80	3.48	.68	2.61	2.72	.11
3.	Use of Resources	2.65	3.20	.55	2.90	3.13	.23	3.04	3.39	.35	2.78	2.32	-.46
4.	Selecting the Plan	2.57	3.22	.65	2.76	3.02	.26	2.81	3.67	.86	2.59	2.41	-.18
5.	Organizing Class	2.74	3.17	.43	2.72	3.08	.36	2.82	3.67	.85	2.96	2.67	-.30
6.	Class Control	2.72	3.04	.31	2.89	3.26	.37	2.89	3.80	.91	2.85	2.69	-.17
7.	Classroom Climate	2.61	3.09	.48	2.70	3.06	.35	2.87	3.76	.89	2.80	2.67	-.14
8.	Student Participation	3.04	3.24	.20	2.89	3.24	.35	3.09	3.95	.85	2.93	2.56	-.37
9.	Measuring Achievement	2.72	2.95	.23	2.92	3.14	.22	2.81	3.52	.70	2.48	2.54	.05
10.	Using Measurement	2.70	2.47	-.23	2.82	2.81	-.02	2.70	3.51	.81	2.47	2.30	-.18
11.	Performances 1 and 2	2.70	3.14	.44	2.88	3.12	.24	2.90	3.51	.61	2.57	2.67	.10
12.	Performances 3, 4 & 5	2.65	3.20	.55	2.79	3.08	.29	2.89	3.57	.68	2.78	2.47	-.31
13.	Performances 6, 7 & 8	2.79	3.12	.33	2.83	3.19	.36	2.95	3.83	.88	2.86	2.64	-.22
14.	Performances 9 and 10	2.72	2.72	.00	2.87	2.98	.11	2.76	3.51	.76	2.48	2.42	-.06
15.	Performances 1 - 10	2.72	3.07	.35	2.84	3.10	.26	2.88	3.63	.75	2.70	2.55	-.15

Figure I

Mean Ratings on All Ten Criteria



of the study, six of the ten behaviors, teaching performances 3 through 8, show the most marked differences between experimental groups and Group IV. These are the behaviors which observers judged most likely to be seen as direct, or overt, behavior by the teacher during the period of observations. On these six behaviors the superiority of the experimental groups to Group IV ranged from 28 per cent to 78 per cent of the mean Group IV ratings.

The analysis of covariance is given in Table 19. This analysis shows no significant difference among the four groups. As in the analysis on the third round, in this final round the total variance among groups was not sufficiently larger than the variance within groups to be considered significant.

Summary of the Data on Classroom Performance

There are no differences in total variance among the four treatment groups which were significant. There were consistent patterns of difference among the groups in certain directions. These patterns are illustrated by the graph in Figure 1 which summarizes these tendencies by showing the means of the four groups calculated from the total judgements of observers on all ten behaviors for each of the four rounds. These consistent patterns of change show:

1. Groups I and II were rated consistently superior to Groups III and IV at the third round of observations.
2. All three experimental groups were superior to Group IV on the final round.
3. Group III was superior to all of the other groups on all behaviors in the final round.

Variations of ratings within treatment groups, the small number of subjects in the final matched group analysis, and the limited range of performance scores all contributed to the finding that variability among the

TABLE 19

ANALYSIS OF COVARIANCE: SIGNIFICANCE OF DIFFERENCES OF MEAN SCORES AMONG THE FOUR GROUPS, EACH OF 18 SUBJECTS, OBTAINED ON THE FOURTH ROUND OF OBSERVATIONS (ROUND 1 OBSERVATION RATINGS USED AS COVARIABLE).

Source	Sums of Squares for Round 4	Sums of Cross-Products	Sums of Squares for Round 1	df
1. Teaching Performance: Suitability of Goals				
Treatments (T)	8.10	4.24	3.72	3
Subjects (S)	31.22	12.30	45.90	17
T x S (Error: E)	110.60	14.14	81.51	51
F ratio = 1.01 (n.s.)				
2. Teaching Performance: Student Acceptance of Goals				
Treatments (T)	5.26	.12	.38	3
Subjects (S)	50.07	1.11	56.49	17
T x S (Error: E)	125.78	- .23	95.42	51
F ratio = .694 (n.s.)				
3. Teaching Performance: Use of Resources				
Treatments (T)	11.97	.78	1.51	3
Subjects (S)	51.09	- 3.86	49.19	17
T x S (Error: E)	130.32	- .59	109.52	51
F ratio = 1.529 (n.s.)				
4. Teaching Performance: Selecting the Plan				
Treatments (T)	14.77	2.12	.78	3
Subjects (S)	30.88	18.61	49.78	17
T x S (Error: E)	134.48	18.28	99.64	51
F ratio = 1.779 (n.s.)				

TABLE 19--Continued

Source	Sums of Squares for Round 4	Sums of Cross-Products	Sums of Squares for Round 1	df
5. Teaching Performance: Organizing the Class				
Treatments (T)	9.10	- 1.20	.65	3
Subjects (S)	32.33	28.95	68.06	17
T x S (Error: E)	146.06	8.11	91.07	51
F ratio = 1.065 (n.s.)				
6. Teaching Performance: Class Control				
Treatments (T)	11.72	.80	.33	3
Subjects (S)	40.79	32.21	68.60	17
T x S (Error: E)	144.89	19.87	81.27	51
F ratio = 1.350 (n.s.)				
7. Teaching Performance: Classroom Climate				
Treatments (T)	11.14	1.08	.70	3
Subjects (S)	36.06	24.92	57.99	17
T x S (Error: E)	135.41	23.38	100.33	51
F ratio = 1.369 (n.s.)				
8. Teaching Performance: Student Participation				
Treatments (T)	17.35	2.09	.49	3
Subjects (S)	34.75	19.96	51.10	17
T x S (Error: E)	127.29	26.55	94.08	51
F ratio = 2.333 (n.s.)				

TABLE 19--Continued

Source	Sums of Squares for Round 4	Sums of Cross-Products	Sums of Squares for Round 1	df
9. Teaching Performance: Measuring Achievement				
Treatments (T)	9.00	3.32	1.87	3
Subjects (S)	35.54	17.54	46.12	17
T x S (Error: E)	108.30	19.42	82.66	51
	F ratio = 1.198 (n.s.)			
10. Teaching Performance: Using Measurement				
Treatments (T)	15.48	2.05	1.16	3
Subjects (S)	52.23	13.88	43.74	17
T x S (Error: E)	106.09	12.32	86.57	51
	F ratio = 2.373 (n.s.)			
11. Step One: Performances 1 and 2				
Treatments (T)	25.78	9.89	5.10	3
Subjects (S)	148.45	71.95	196.93	17
T x S (Error: E)	445.01	71.59	335.98	51
	F ratio = .836 (n.s.)			
12. Step Two: Performances 3, 4 and 5				
Treatments (T)	102.82	6.60	4.51	3
Subjects (S)	292.25	194.19	484.23	17
T x S (Error: E)	1168.08	102.35	848.68	51
	F ratio = 1.460 (n.s.)			

TABLE 19--Continued

Source	Sums of Squares for Round 4	Sums of Cross-Products	Sums of Squares for Round 1	df
13. Step Three: Performances 6, 7 and 8				
Treatments (T)	117.57	10.15	2.29	3
Subjects (S)	315.48	225.49	509.53	17
T x S (Error: E)	1162.96	198.43	758.12	51
F ratio = 1.685 (n.s.)				
14. Step Four: Performances 9 and 10				
Treatments (T)	46.83	11.07	5.92	3
Subjects (S)	166.58	63.41	178.13	17
T x S (Error: E)	411.55	62.33	331.50	51
F ratio = 1.774 (n.s.)				
15. All Performances (total)				
Treatments (T)	1050.55	179.78	42.92	3
Subjects (S)	3441.69	2051.80	5023.95	17
T x S (Error: E)	1631.73	1662.04	8152.93	51
F ratio = 1.440 (n.s.)				

groups was not significant. The consistent pattern of differences, however, does not appear to the staff to be a finding which can be attributed to chance, and is thus considered to be a significant result of this analysis.

Analysis of the Data on Teacher Attitudes

The revised attitude scale was administered to all the beginning teachers in the study during each round of observations. The scale consisted of seven concepts related to the profession of teaching which appeared to discriminate among subjects during the pilot study. These concepts were tests, teachers' meetings, school administrators, homework, teaching work load, grading and teaching as a career. Adjective scales were taken from Osgood's thesaurus analysis. These scales were tested during the pilot study, and those scales which appeared to be most promising as discriminators were utilized. Particular adjective scales varied somewhat from concept to concept, depending upon the evidence obtained in the pilot study. The revised scale appears in Appendix B.

The first administration of the attitude scale, during the first round of observations, took place between the second and fourth weeks of school. Each subject was mailed a copy of the attitude test before the observation team's first visit. At the conclusion of the observation the team picked up the completed attitude test from the subject. The same procedure was followed in the later administration of the test.

Mean scores and standard deviations for each treatment group of beginning teachers for each adjective pair on each concept are reported in Table 28 in Appendix B. The same table shows these data for all four administrations of this semantic differential attitude scale. Low scores indicate relatively high or favorable attitudes, and high scores indicate low or unfavorable attitudes.

The initial assessment of attitudes indicated a generally favorable feeling towards the profession of teaching. On the second administration of the attitude test, four and a half weeks after the first administration, most of the mean scores of the groups were in the negative direction. (Table 20) There was not much net change in attitude by the entire population of the study after the second administration. A possible exception was the changes in attitudes reflected towards the concept, "Teaching Work Load". The beginning teachers showed a strong negative change towards this concept on the second round of appraisals, followed by a strong positive change on the third round of administrations.

Table 21 reports the net changes in attitude from the first to the last administration of the test by treatment groups. Only members of Group II changed their responses to the attitude scale in a positive direction. Group II attitude scores were slightly lower than other groups at the beginning of the study, which may account partly for this difference in pattern of attitude changes.

Table 22 shows the analysis of variance of group mean scores on the attitude scale when data for all four rounds are grouped. Significant differences between rounds, or administrations, were found for two of the seven concepts. There were no significant differences among treatment groups. The significant F ratios between rounds, together with results of simple inspection of the direction of attitude changes, suggest that there was a significant tendency for all the beginning teachers to show more negative attitudes toward teaching after the first month of experience.

The F ratios obtained for each round in the analysis of variance among

TABLE 20

DIRECTION OF ATTITUDE CHANGES BY TREATMENT GROUPS SHOWN ON FOUR ADMINISTRATIONS OF THE SEMANTIC DIFFERENTIAL SCALE

Concept	From Round 1 to 2[a]		From Round 2 to 3[b]		From Round 3 to 4[c]	
	No. Positive	No. Negative	No Positive	No Negative	No. Positive	No. Negative
1. Tests	6	24	13	17	16	16
2. Teachers' Meetings	4	28	15	17	17	15
3. School Administrators	9	19	7	21	14	13
4. Homework	8	22	18	14	18	13
5. Teaching Work Load	3	23	20	8	17	10
6. Grading	11	14	18	9	4	16
7. Teaching as a Career	6	18	8	16	12	11

[a] No changes 9 comparisons [b] No changes 3 comparisons [c] No changes 4 comparisons

TABLE 21

NET CHANGES IN ATTITUDES FROM THE BEGINNING TO THE END OF THE STUDY BY THE FOUR TREATMENT GROUPS

Group	Direction of Change of Attitude on Adjective Pairs		
	Negative	Positive	No Change
Group I	39	11	1
Group II	24	27	-
Group III	43	7	1
Group IV	39	9	3

TABLE 22

ANALYSIS OF VARIANCE AMONG SUBJECTS TESTED WITHIN FOUR TREATMENT GROUPS ON FOUR ROUNDS OF THE SEMATIC DIFFERENTIAL ATTITUDE TEST

Source of Variance	Sums of Squares	Mean Squares	df	F ratios
1. Concept: Tests				
Treatments (T)	294.80	98.27	3	1.10^{a}
Rounds (R)	76.80	25.60	3	1.99^{b}
Subjects tested within Treatments (S:T)	8,208.90	89.23	92	
Treatments x Rounds (TR)	117.20	13.02	9	1.01^{c}
Rounds x Subjects (R S:T)	3,558.20	12.89	276	
2. Concept: Teacher's Meetings				
Treatments (T)	520.20	173.40	3	.95^{a}
Rounds (R)	188.70	62.90	3	3.60^{b}*
Subjects tested within Treatments (S:T)	16,704.90	181.58	92	
Treatments x Rounds (TR)	112.70	12.52	9	.72^{c}
Rounds x Subjects (R S:T)	4,820.80	17.47	276	
3. Concept: School Administrators				
Treatments (T)	339.90	113.30	3	1.47^{a}
Rounds (R)	130.90	43.63	3	3.93^{b}*
Subjects tested within Treatments (S:T)	7,099.60	77.17	92	
Treatments x Rounds (TR)	93.90	10.43	9	.94^{c}
Rounds x Subjects (R S:T)	3,063.50	11.10	276	

[a] $F_T = \dfrac{MS_T}{MS_{S:T}}$

[b] $F_R = \dfrac{MS_R}{MS_{RS:T}}$

[c] $F_T = \dfrac{MS_{TR}}{MS_{R:ST}}$

* Significant at .05 level

TABLE 22-Continued

Source of Variance	Sums of Squares	Mean Squares	df	F ratios
4. Concept: Homework				
Treatments (T)	232.90	77.63	3	.42[a]
Rounds (R)	73.20	24.40	3	.32[b]
Subjects tested within Treatments (S:T)	17,095.80	185.82	92	
Treatments x Rounds (TR)	69.20	7.69	9	.42[c]
Rounds x Subjects (R S:T)	5,099.98	18.48	276	
5. Concept: Teaching Work Load				
Treatments (T)	567.60	189.20	3	1.30[a]
Rounds (R)	70.60	23.53	3	1.81[b]
Subjects tested within Treatments (S:T)	13,366.60	145.29	92	
Treatments x Rounds (TR)	87.60	9.73	9	.75[c]
Rounds x Subjects (R S:T)	3,597.20	13.03	276	
6. Concept: Grading				
Treatments (T)	607.00	202.33	3	1.10[a]
Rounds (R)	63.10	21.03	3	.96[b]
Subjects tested within Treatments (S:T)	16,877.20	183.45	92	
Treatments x Rounds (TR)	256.90	28.54	9	1.30[c]
Rounds x Subjects (R S:T)	6,054.20	21.94	276	
7. Concept: Teahing as a Career				
Treatments (T)	132.90	44.30	3	.97[a]
Rounds (R)	34.50	11.50	3	2.26[b]
Subjects tested within Treatments (S:T)	4,181.00	45.45	92	
Treatments x Rounds (TR)	63.90	7.10	9	1.39[c]
Rounds x Subjects (R S:T)	1,405.00	5.09	276	

the four treatment groups, is shown in Table 23. This table seems to show that differences among the four groups as to attitude were slightly more pronounced, not significantly, at the beginning of the study and less so at the end of the study.

Inspection of Table 28, Appendix B, shows that there was little net change in attitude during the period of the study. Only 15 of 204 group mean scores on different adjective pairs of the seven concepts tested changed by as much as half of a point on the seven-point scale used in the semantic differential attitude test.

The data obtained from the semantic differential attitude scale do not appear to show significant differences which are related to the experimental treatments of the subjects in this study.

Characteristics of Teachers Related to Performance

Two characteristics of the subjects in the study were correlated with teaching performance. These characteristics were cumulative grade point average earned in the bachelor degree program and the grade point average of courses in Education. These two averages were correlated with the mean ratings of observer teams on each of the ten teaching behaviors evaluated in the study. Table 24 reports these correlations for the final round of observations. Both sets of correlations show slightly negative relationships, although for practical purposes these correlations approach zero. There appears to be no correlation of grade point earned in college with performance as rated by trained observer teams. Other data regarding characteristics of beginning teachers which might have been related to performance in this study resulted in such small sub-groups, that no practical purpose seemed likely to be served by further analysis.

Summary: The Evidence as to the Major Hypotheses

Hypothesis 1 was that there would be no differences in selected aspects of classroom behavior between groups of elementary teachers who have a reduction in load compared with others with no reduction. All three of the experimental groups were given a reduction in load. Compared to Group IV teachers, Group III teachers had 25 per cent fewer pupils and Groups I and II teachers met their pupils for 25 per cent less time during each day.

The analysis of covariance supports the hypothesis in that no significant differences among groups were obtained. The observed patterns of differences would reject the hypothesis in that 28 out of 30 comparisons of Groups I, II and III teachers with Group IV on the final round of appraisals of classroom performance showed that each of the experimental groups were superior to the control group.

The research staff believed that the consistent findings showing superior classroom performances by teachers with reduced loads, together with the relatively large percentage differences between mean scores, can not be explained as being likely to result from pure chance. The staff judged that the weight of evidence would reject the hypothesis, but that others may judge differently from these same data.

Hypothesis 2 was that there will be no significant differences in attitude between groups of elementary teachers with reduced loads and those with no reduced loads. Again the comparison to test the hypothesis was between Groups I, II and III with Group IV. The analysis of variance produced no significant differences, neither were there observed patterns of differences. Regardless of load, beginning teachers tend to change their attitudes towards teaching in a negative direction during the first five months of school. Hypothesis 2 was not rejected by the evidence.

TABLE 23

F RATIOS FROM ANALYSES OF VARIANCE AMONG TREATMENT GROUPS ON EACH OF FOUR ADMINISTRATIONS OF THE SEMANTIC DIFFERENTIAL ATTITUDE TEST
(N-92 F of 2.72 required for .05 level of significance)

Concept	F Ratios			
	Round 1	Round 2	Round 3	Round 4
1. Tests	.56	.54	2.23	1.06
2. Teachers' Meetings	1.77	.91	.62	.47
3. Administrators	2.71	1.38	1.07	.51
4. Homework	.17	.73	.24	.48
5. Teaching Work Load	2.28	.69	.79	1.20
6. Grading	2.12	.65	.86	.99
7. Teaching as a Career	2.51	1.16	.92	.37

TABLE 24

CORRELATION COEFFICIENTS OF ALL-COLLEGE GRADE POINT AVERAGES AND EDUCATION COURSE GPA WITH FINAL RATINGS OF TEACHING PERFORMANCES

Teaching Performance	r with College GPA	r with Education Course GPA
1. Suitability of Goals	-.12	-.09
2. Student Acceptance of Goals	-.13	-.07
3. Use of Resources	-.11	-.03
4. Selecting the Plan	-.11	-.08
5. Organizing the Class	-.15	-.07
6. Classroom Control	-.16	-.07
7. Classroom Climate	-.15	-.04
8. Student Participation	-.12	-.03
9. Measuring Achievement	-.16	-.10
10. Using Measurement	-.13	-.08
Total, 1. - 10.	-.14	-.07

Hypothesis 3 was that there will be no significant differences in terms of selected aspects of classroom behavior between groups of beginning elementary teachers who received different types of inservice instruction during released time periods. To test this hypothesis, Groups I and II were compared. Neither the analysis of covariance or the inspection of patterns of differences produced evidence which would reject this hypothesis.

Hypothesis 4 was that there will be no differences in attitudes between groups of teachers who received different types of inservice instruction during released time periods. Analysis of variance showed no differences in the semantic differential attitude scale between Groups I and II. Group II differed somewhat from Group I (and also from Groups III and IV) in that teachers of Group II showed more evidence of change in attitude in a positive direction during the period of the study. The net changes in attitude, positive or negative, did not appear to be large, and the research staff judged that the evidence of the study supported, rather than rejected, Hypothesis 4.

CHAPTER VI

SUMMARY, CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

This study was an attempt to see whether experimental treatments involving reduced loads and intensive inservice instruction would affect the performances and attitudes of beginning teachers. The treatments were chosen because an advisory committee to the research project believed in the efficacy of these treatments on the bases of their respective backgrounds of experience. The 25 per cent reduction in load and the inservice instruction which was given twice a week over a period of one semester as treatments to experimental groups of teachers consituted rather gross variables. It was hoped that these treatments would be sufficiently powerful to be reflected in changes in behavior, but the treatments were deliberately kept to a moderate degree of intensity in hopes that the project staff could make judgements as to the feasibility of administering these treatments. Questions as to cost of the treatments and the degree to which the education of teachers would be prolonged were factors considered subjectively in setting the intensity of the experimental treatments.

The project staff hoped to conclude from the data of the study to what extent an internship program for career teachers would be justified. The practical purpose for the study, then, was to determine the degree to which internship programs, based upon the experimental treatments in the project, would bring about changes in behavior which were proportionate to the cost and effort expended. The staff also hoped to be able to identify questions which would be important to study with more refined research designs.

The measures of performance and attitude used were chosen because they seemed likely to reveal changes in behavior and would yield reliable data for comparing subjects. The purpose of the instruments used was to identify differences among subjects, rather than to help them improve their teaching.

The observation check list is not claimed to be a valid way of assessing the entire teaching act, but was chosen because it seemed a promising tool for appraising a sample of a teacher's classroom skills. The semantic differential attitude scale was selected because it was easily administered and other studies suggested it could be sensitive to differences among subjects on concepts which seem to reflect attitudes. Finally, the evaluation instruments were chosen partly because different members of the advisory committee had experience in other studies with these instruments.

The experience in this study with the appraisal techniques used to evaluate teachers' performances suggests that experienced teachers can be trained to use an observation check list and arrive at a fair degree of agreement with other such trained observers. Judging from the ratings of groups of teachers by teams of observers, the trained observer tends to see different behaviors defined in the check-list as different ways of judging one, general teaching performance. Also, two of the observers in the study rated teachers somewhat differently than the other seven. A limitation to this appraisal procedure is the possibility that some observers will differ from the majority depending upon personal characteristics which are not known to the staff of this project.

A possible subject for further study would be the extent to which different kinds and amounts of professional experience affect observers' appraisals of teaching performance.

The observers in this project were trained in seminar sessions with Dr. Harry Garrison, who helped develop the Stanford Teacher Competence Appraisal Guide and who adapted the Stanford quide in the form used for this project. The observers then practiced using the checklist in observations of experienced, able Seattle teachers and in observations of student teachers from colleges in the Seattle area. In this way they were able to see a range of teaching behavior. They also practiced by using the checklist to evaluate teaching behavior which had been video-taped. In this way they were able to test their judgements by comparing them with other observers.

At the start of the project, the observers rated the beginning teachers somewhat below what they conceived to be average for teachers in general. After about a month, the observers rated these teachers a second time and found their performances to be lower than on the first round. These lower ratings were reported for all teachers regardless of treatment. The group of beginning teachers who received no special treatment remained at this low level of rated performance throughout the period of the study. Two of the three groups receiving treatments showed marked improvement on the third round of appraisals. The experimental group which showed no improvement on the third round was rated the highest of all on the fourth and final round. At the conclusion of the study all three treatment groups were considered superior to the no-treatment group on 28 of 30 comparisons. The differences between means of each of the three experimental groups and the control group on total teaching performance were between 26 and 58 per cent of the control group's mean score.

In spite of the large and consistent apparent differences between groups, analysis of covariance showed that these differences were not significantly greater than the variance within groups. One reason for this was the shrinkage in the size of the sample. Subjects were assigned to treatments randomly after being matched in sets of four subjects. When one subject was lost because of illness or some other reason, the project lost that subject's counterparts in each of the other groups. By the end of the study for analysis purposes there were 18 subjects in each group instead of the 26 in each group who actually were appraised in the first round of observations. Another reason for the lack of significance in the analysis of covariance was the variation in observer ratings within the different teams. Some agreement was obtained but the correlations between observers were lower than hoped. Another reason for lack of significance in the analysis of covariance was that two of the experimental groups tended to achieve mean ratings which were almost the same. Thus variability among all four groups was limited by the fact that experimental groups were not different from each other, even though there were large apparent differences between these groups and the control group.

Although all three of the experimental groups were rated superior to the control group by the end of the study, there were apparent differences among the three experimental groups. Two of the experimental groups received intensive inservice help in addition to reduced leads. The third group received only reduced loads in the form of fewer pupils. At the third round the groups with inservice treatments appeared to be markedly superior to both the control group and the other experimental group. At the final round, however, the group with only reduced loads as treatments was about the same amount superior to the two inservice groups as these two treatment groups were superior to the control group. Why Group III, the small-class group, made a sudden improvement in teaching performance is not clear.

Bias on the part of the observer teams was considered as a possible explanation for the high final Group III scores. Although the observers were not told to which group any of the beginning teachers were assigned, they could have identified Group III subjects by noting the relatively small numbers of pupils in these teachers' classes. If this knowledge could have been reflected in biased ratings, it is still difficult to reconcile the fact that observer teams consistently rated Group III teachers low for three rounds and then high on the last round. There seems no reason for a change in bias, if such existed.

The pattern of improvement in teaching performance on the part of the treatment groups receiving both inservice help and reduced loads suggests the major change resulting from this treatment occured in the first four months and little change occurred after that. (Figure 1) This pattern suggests an intensive internship type of treatment for beginning teachers during the fall of the first year of experience. The Group III pattern, however, is inconclusive as to the amount of time necessary for intensive treatment. From the evidence in this study, we have no way of predicting the level of teaching behavior on the part of Group III after January. A longer study would be necessary to determine if the apparent high levels of performances of Group III teachers would be sustained. Why having small classes was not effective in raising levels of teaching performance for four months, and suddenly became effective is not clear. There seems no explanation, however, other than the change in behavior by Group III teachers was related to small classes. Inspection of the patterns of changed teaching behavior on the part of the experimental groups suggests that an internship program should be extended for at least a semester and probably for the first year. The amounts of intensive treatment in the form of inservice instruction, however, might be reduced in the second semester.

The attitude changes of the teachers in this study did not appear to be of very great magnitude. There were significant changes in a negative direction after the first month of school. All the treatment groups indicated a more negative attitude towards school-related concepts on the second administration of the attitude scale than on the first administration. This lower level of attitudes was maintained throughout the remainder of the study. Changes in attitudes, however, did not appear to be related to treatments.

Success of the teacher in courses while at college does not appear to be related to teaching performance. This seems to be true both of the over-all academic performance of the teacher while at college and those courses which were part of the professional education requirement. Correlations of grade point and teaching performance as rated by the observer teams were consistently negative, but not to a significant level.

In the judgement of the staff the change in classroom performance of the experimental groups and the lack of change on the part of the control group justify the development of a plan for internship as training for career teachers in the elementary school. In spite of the gross character of the experimental variables and the difficulty of controlling all factors which might have affected the results of the study, the general findings about teachers' behavior were such that the staff feels that it is highly unlikely that variables other than the treatments administered to the experimental groups could have accounted for the observed differences.

Questions for Further Study

1. Will the advantage in teaching skill apparently demonstreated by teachers in the experimental groups of this study be observeable at the end of one academic year; at some point during the second year of teaching?

2. Will a combination of small class size (Group III treatment) and intensive inservice instruction of a type comparable to that administered to Groups I and II contribute to a level of teaching performance higher than that achieved by teachers given either of the three treatments in this project?

3. Will the general conclusions of this study as to the effects of reduced load and inservice instruction apply to beginning secondary teachers

as well as elementary teachers?

4. Would recommended treatments other than those administered to the subjects of this study contribute to similar or different levels of teaching performance?

5. Will beginning teachers given an internship experience similar to the experimental treatments in this study remain in teaching longer than the average beginning teacher, or otherwise demonstrate a high level of committment to the profession of teaching?

6. What special competencies and experiences are required for the person who provides intensive inservice instruction to the beginning teacher?

7. If future teachers while in college are given systematic instruction and practice in demonstrating the specific behaviors evaluated in this study, will they then demonstrate these behaviors at a significantly higher level than comparable beginning teachers?

BIBLIOGRAPHY

Anderson, H. H., and Brewer, Helen M. "Studies of Teachers' Classroom Personalities II: Effects of Teacher's Dominative and Integrative Contacts on Children's Classroom Behavior", Applied Psychological Monographs, 1946, No. 8.

Barr, A. S., "The Measurement and Prediction of Teaching Efficiency: A Summary of Investigations", Journal of Experimental Education, XVI (June, 1948), 203-283.

Barr, A. S. et al."Wisconsin Studies of the Measurement of Teaching Effectiveness," Madison, Wisconsin: Dunbar Publications, Inc., 1961

Biddle, J. B., and Ellena, William J. "Contemporary Research on Teacher Effectiveness." Chicago: Holt, Rinehart, and Winston, 1964

Brain, George. "Evaluating Teacher Effectiveness", NEA Journal, 54 (February, 1965), 35-36

Conant, James B. "The Education of the American Teacher." McGraw-Hill Co., 1963

Domas, Simeon J., and Tiedeman, D. V. "Teacher Competence: An Annotated Bibliography", Journal of Experimental Education, XIX (December, 1950), 101-218

Flanders, Ned. "Teacher Influence, Pupil Attitudes and Achievement: Studies in Interaction Analysis". Minneapolis: University of Minnesota, U. S. Office of Education Cooperative Research Project 397, 1960. (Mimeographed)

________."Interaction Analysis: A Technique for Quantifying Teacher Influence", Research Monograph BER (Minneapolis, Minn.: University of Minnesota) 61-2, 1961

Gage, N. L. (ed.) "Handbook of Research on Teaching". Chicago: Rand McNally and Co., 1963

Garrett, Henry. "Statistics in Psychology and Education", New York: David McKay Co., Inc., 1962

Garrison, Harry L., "Evaluation of Teaching and Learning",(unpublished Ed.D. thesis, Stanford University, 1964)

Getzels, Jacob, and Thelen, Herbert. "The Classroom Group As a Unique Social System", 59th Yearbook, N.S.S.E., 1962, 53-82

Howsam, Robert. "Who's a Good Teacher?" Burlingame, California: Joint Committee on Personnel Procedures, CSBA and CTA, 1960

Hughes, Marie. "Patterns of Effective Teaching, Second Progress Report, Merit Study". Provo, Utah: Provo City Schools, 1961

"Instructions for Users: The Seattle Teaching Performance Appraisal Manual." Seattle, Washington: Seattle Public Schools, Experimental, 1965-66

Jersild, A. J. et al. "An Evaluation of Aspects of the Activity Program in the New York City Public Elementary Schools", Journal of Experimental Education, 8 (August, 1939), 166-207.

Lindquist, E. F. "Design and Analysis of Experiments in Psychology and Education", Houghton Mifflin Co., Boston 1956.

Lindsay, Margaret. "New Horizons for the Teaching Profession". Washington, D. C.: NEA TEPS Commission, 1961.

"Manual to Assist Staff Using the Stanford Appraisal Guides", Rev. ed. Secondary Education Project. Stanford: Stanford University Press, 1963

McCall, William A. "Measurement of Teacher Merit", Raleigh, N. C.: State Department of Education, 1962

Medley, Donald M., and Mitzel, Harold E. "Studies of Teacher Behavior: Refinement of Two Techniques for Assessing Teacher's Classroom Behaviors" New York: Board of Higher Education, City of New York, Division of Teacher Education, Office of Research and Evaluation, 1955.

Morsh, J. E. "Development Report - Systematic Observation of Instructor Behavior", USAF Personnel Training Research Center Development Report, 1956 No. AFPTRC-TN-56-52.

National Education Association, National Commission on Teacher Education and Professional Standards. "The Real World of the Beginning Teacher". Report of the Nineteenth National TEPS Conference. Washington D. C.: The Commission, 1966. 90 pp.

Osgood, Charles E., Suci, George J., and Tannenbaum, Percy H., "The Measurement of Meaning", Urbana, Illinois: University of Illinois Press, 1957.

Smith, B. O. and Ennis, R. H. "Language and Concepts of Education", New York: Rand McNally, Inc., 1961.

APPENDIX A

FORMS USED IN THE STUDY

RYANS' CLASSROOM OBSERVATION RECORD

Teacher ____________________ Sex ____ Class or Subject __________ Date __________

Observation No. ________ School ______________ Time ______ Observer __________

PUPIL BEHAVIOR			REMARKS
1. Apathetic	1 2 3 4 5 6 7 N	Alert	________
2. Obstructive	1 2 3 4 5 6 7 N	Responsible	________
3. Uncertain	1 2 3 4 5 6 7 N	Confident	________
4. Dependent	1 2 3 4 5 6 7 N	Initiating	________
TEACHER BEHAVIOR			
5. Partial	1 2 3 4 5 6 7 N	Fair	________
6. Autocratic	1 2 3 4 5 6 7 N	Democratic	________
7. Aloof	1 2 3 4 5 6 7 N	Responsive	________
8. Restricted	1 2 3 4 5 6 7 N	Understanding	________
9. Harsh	1 2 3 4 5 6 7 N	Kindly	________
10. Dull	1 2 3 4 5 6 7 N	Stimulating	________
11. Stereotyped	1 2 3 4 5 6 7 N	Original	________
12. Apathetic	1 2 3 4 5 6 7 N	Alert	________
13. Unimpressive	1 2 3 4 5 6 7 N	Attractive	________
14. Evading	1 2 3 4 5 6 7 N	Responsible	________
15. Erratic	1 2 3 4 5 6 7 N	Steady	________
16. Excitable	1 2 3 4 5 6 7 N	Poised	________
17. Uncertain	1 2 3 4 5 6 7 N	Confident	________
18. Disorganized	1 2 3 4 5 6 7 N	Systematic	________
19. Inflexible	1 2 3 4 5 6 7 N	Adaptable	________
20. Pessimistic	1 2 3 4 5 6 7 N	Optimistic	________
21. Immature	1 2 3 4 5 6 7 N	Integrated	________
22. Narrow	1 2 3 4 5 6 7 N	Broad	________

RYANS' CLASSROOM OBSERVATION RECORD

Teacher__________________________ Sex____ Class or Subject_________ Date_________

Observation No.________ School_____________ Time_____ Observer___________

PUPIL BEHAVIOR				REMARKS
1.	**Apathetic**	**1 2 3 4 5 6 7 N**	**Alert**	____________
2.	**Obstructive**	**1 2 3 4 5 6 7 N**	**Responsible**	____________
3.	**Uncertain**	**1 2 3 4 5 6 7 N**	**Confident**	____________
4.	**Dependent**	**1 2 3 4 5 6 7 N**	**Initiating**	____________
TEACHER BEHAVIOR				
5.	**Partial**	**1 2 3 4 5 6 7 N**	**Fair**	____________
6.	**Autocratic**	**1 2 3 4 5 6 7 N**	**Democratic**	____________
7.	**Aloof**	**1 2 3 4 5 6 7 N**	**Responsive**	____________
8.	**Restricted**	**1 2 3 4 5 6 7 N**	**Understanding**	____________
9.	**Harsh**	**1 2 3 4 5 6 7 N**	**Kindly**	____________
10.	**Dull**	**1 2 3 4 5 6 7 N**	**Stimulating**	____________
11.	**Stereotyped**	**1 2 3 4 5 6 7 N**	**Original**	____________
12.	**Apathetic**	**1 2 3 4 5 6 7 N**	**Alert**	____________
13.	**Unimpressive**	**1 2 3 4 5 6 7 N**	**Attractive**	____________
14.	**Evading**	**1 2 3 4 5 6 7 N**	**Responsible**	____________
15.	**Erratic**	**1 2 3 4 5 6 7 N**	**Steady**	____________
16.	**Excitable**	**1 2 3 4 5 6 7 N**	**Poised**	____________
17.	**Uncertain**	**1 2 3 4 5 6 7 N**	**Confident**	____________
18.	**Disorganized**	**1 2 3 4 5 6 7 N**	**Systematic**	____________
19.	**Inflexible**	**1 2 3 4 5 6 7 N**	**Adaptable**	____________
20.	**Pessimistic**	**1 2 3 4 5 6 7 N**	**Optimistic**	____________
21.	**Immature**	**1 2 3 4 5 6 7 N**	**Integrated**	____________
22.	**Narrow**	**1 2 3 4 5 6 7 N**	**Broad**	____________

GLOSSARY

(To be used with Ryans' Classroom Observation Record)

PUPIL BEHAVIORS

1. Apathetic-Alert Pupil Behavior

Apathetic	Alert
1. Listless	1. Appeared anxious to recite and participate
2. Bored-acting	2. Watched teacher attentively
3. Entered into activities half-heartedly	3. Worked concentratedly
4. Restless	4. Seemed to respond eagerly
5. Attention wandered	5. Prompt and ready to take part in activities when they begin
6. Slow in getting under way	

2. Obstructive-Responsible Pupil Behavior

Obstructive	Responsible
1. Rude to one another and/or to teacher	1. Courteous, cooperative, friendly with each other and with teacher
2. Interrupting; demanding attention; disturbing	2. Completed assignments without complaining or unhappiness
3. Obstinate; sullen	3. Controlled voices
4. Refusal to participate	4. Received help and criticism attentively
5. Quarrelsome; irritable	5. Asked for help when needed
6. Engaged in name-calling and/or tattling	6. Orderly without specific directions from teacher
7. Unprepared	7. Prepared

3. Uncertain-Confident Pupil Behavior

Uncertain	Confident
1. Seemed afraid to try; unsure	1. Seemed anxious to try new problems or activities
2. Hesitant; restrained	2. Undisturbed by mistakes
3. Appeared embarrassed	3. Volunteered to recite
4. Frequent display of nervous habits, nail-biting, etc.	4. Entered freely into activities
5. Appeared shy and timid	5. Appeared relaxed
6. Hesitant and/or stammering speech	6. Spoke with assurance

4. Dependent-Initiating Pupil Behavior

Dependent	Initiating
1. Relied on teacher for explicit directions	1. Volunteered ideas and suggestions
2. Showed little ability to work things out for selves	2. Showed resourcefulness
3. Unable to proceed when initiative called for	3. Took lead willingly
4. Appeared reluctant to take lead or to accept responsibility	4. Assumed responsibilities without evasion

TEACHER BEHAVIORS

5. Partial-Fair Teavher Behavior

Partial	Fair
1. Repeatedly slighted a pupil	1. Treated all pupils approximately equally
2. Corrected or criticized certain pupils repeatedly	2. In case of controversy pupil allowed to explain his side
3. Repeatedly gave a pupil special advantages	3. Distributed attention to many pupils
4. Gave most attention to one or a few pupils	4. Rotated leadership impartially
5. Showed prejudice (favorable or unfavorable) toward some social, racial, or religious groups	5. Based criticism or praise on factual evidence, not hearsay
6. Expressed suspicion of motives of a pupil	

6. Autocratic-Democratic Teacher Behavior

Autocratic	Democratic
1. Told pupils each step to take	1. Guided pupils without being mandatory
2. Intolerant of pupils' ideas	2. Exchanged ideas with pupils
3. Mandatory in giving directions; orders to be obeyed at once	3. Encouraged (asked for) pupil opinion
4. Interrupted pupils although their discussion was relevant	4. Encouraged upils to make own decisions
5. Always directed rather than participated	5. Entered into activities without domination

7. Aloof-Responsive Teacher Behavior

Aloof	Responsive
1. Stiff and formal in relations with pupils	1. Approachable to all pupils

2. Apart; removed from class activity	2. Participated in class activity
3. Condescending to pupils	3. Responded to reasonable requests and/or questions
4. Routine and subject matter only concern; pupils as persons ignored	4. Spoke to pupils as equals
5. Referred to pupil as "this child" or "that child"	5. Commended effort
	6. Gave encouragement
	7. Recognized individual differences

8. Restricted-Understanding Teacher Behavior

Restricted	Understanding
1. Recognized only academic accomplishments of pupils; no concern for personal problems	1. Showed awareness of a pupil's personal emotional problems and needs
2. Completely unsympathetic with a pupil's failure at a task	2. Was tolerant of error on part of pupil
3. Called attention only to very good or very poor work	3. Patient with a pupil beyond ordinary limits of patience
4. Was impatient with a pupil	4. Showed what appeared to be sincere sympathy with a pupil's viewpoint

9. Harsh-Kindly Teacher Behavior

Harsh	Kindly
1. Hypercritical; fault-finding	1. Went out of way to be pleasant and/or to help pupils; friendly
2. Cross; curt	2. Gave a pupil a deserved compliment
3. Depreciated pupil's efforts; was sarcastic	3. Found good things in pupils to call attention to
4. Scolded a great deal	4. Seemed to show sincere concern for a pupil's personal problem
5. Lost temper	5. Showed affection without being demonstrative
6. Used threats	6. Disengaged self from a pupil without bluntness
7. Permitted pupils to laugh at mistakes of others	

10. Dull-Stimulating Teacher Behavior

Dull	Stimulating
1. Uninteresting, monotonous explanations	1. Highly interesting presentation; got and held attention without being flashy
2. Assignments provided little or no motivation	2. Clever and witty, though not smart-alecky or wisecracking
3. Failed to provide challenge	3. Enthusiastic; animated
4. Lacked animation	4. Assignments challenging
5. Failed to capitalize on pupil interests	5. Took advantage of pupil interests
6. Pedantic, boring	6. Brought lesson successfully to a climax
7. Lacked enthusiasm; bored-acting	7. Seemed to provoke thinking

11. Stereotyped-Original Teacher Behavior

Stereotyped

1. Used routine procedures without variation
2. Would not depart from procedure to take advantage of a relevant question or situation
3. Presentation seemed unimaginative
4. Not resourceful in answering questions or providing explanations

Original

1. Used what seemed to be original and relatively unique devices to aid instruction
2. Tried new materials or methods
3. Seemed imaginative and able to develop presentation around a question or situation
4. Resourceful in answering questions; had many pertinent illustrations available

12. Apathetic-Alert Teacher Behavior

Apathetic

1. Seemed listless, languid, lacked enthusiasm
2. Seemed bored by pupils
3. Passive in response to pupils
4. Seemed preoccupied
5. Attention seemed to wander
6. Sat in chair most of time; took no active part in class activities

Alert

1. Appeared buoyant; wide-awake; enthusiastic about activity of the moment
2. Kept constructively busy
3. Gave attention to, and seemed interested in, what was going on in class
4. Prompt to "pick up" class when pupils' attention showed signs of lagging

13. Unimpressive-Attractive Teacher Behavior

Unimpressive

1. Untidy or sloppily dressed
2. Inappropriately dressed
3. Drab, colorless
4. Posture and bearing unattractive
5. Possessed distracting personal habits
6. Mumbled; inaudible speech; limited expression; disagreeable voice tone; poor inflection

Attractive

1. Clean and neat
2. Well-groomed; dress showed good taste
3. Posture and bearing attractive
4. Free from distracting personal habits
5. Plainly audible speech; good expression; agreeable voice tone; good inflection

14. Evading-Responsible Teacher Behavior

Evading

1. Avoided responsibility; disinclined to make decisions

Responsible

1. Assumed responsibility; made decisions as required

2. "Passed the buck" to class, to other teachers, etc.	2. Conscientious
3. Left learning to pupil, failing to give adequate help	3. Punctual
4. Let a difficult situation get out of control	4. Painstaking; careful
5. Assignments and directions indefinite	5. Suggested aids to learning
6. No insistence on either individual or group standards	6. Controlled a difficult situation
7. Inattentive with pupils	7. Gave definite directions
8. Cursory	8. Called attention to standards of quality
	9. Attentive to class
	10. Thorough

15. Erratic-Steady Teacher Behavior

Erratic	Steady
1. Impulsive; uncontrolled; temperamental; unsteady	1. Calm; controlled
2. Course of action easily swayed by circumstances of the moment	2. Maintained progress toward objective
3. Inconsistent	3. Stable, consistent, predictable

16. Excitable-Poised Teacher Behavior

Excitable	Poised
1. Easily distrubed and upset; flustered by classroom situation	1. Seemed at ease at all times
2. Hurried in class activities; spoke rapidly using many words and gestures	2. Unruffled by situation that developed in classroom; dignified without being stiff or formal
3. Was "jumpy"; nervous	3. Unhurried in class activities; spoke quietly and slowly
	4. Successfully diverted attention from a stress situation in classroom

17. Uncertain-Confident Teacher Behavior

Uncertain	Confident
1. Seemed unsure of self; faltering, hesitant	1. Seemed sure of self; self-confident in relations with pupils
2. Appeared timid and shy	2. Undisturbed and unembarrassed by mistakes and/or criticism
3. Appeared artificial	

18. Disorganized-Systematic Teacher Behavior

Disorganized

1. No plan for classwork
2. Unprepared
3. Objectives not apparent; undecided as to next step
4. Wasted time
5. Explanations not to the point
6. Easily distracted from matter at hand

Systematic

1. Evidence of a planned though flexible procedure
2. Well prepared
3. Careful in planning with pupils
4. Systematic about procedure of class
5. Had anticipated needs
6. Provided reasonable explanations
7. Held discussion together; objectives apparent

19. Inflexible-Adaptable Teacher Behavior

Inflexible

1. Rigid in conforming to routine
2. Made no attempt to adapt materials to individual pupils
3. Appeared incapable of modifying explanation or activities to meet particular classroom situations
4. Impatient with interruptions and digressions

Adaptable

1. Flexible in adapting explanations
2. Individualized materials for pupils as required; adapted activities to pupils
3. Took advantage of pupils' questions to further clarify ideas
4. Met an unusual classroom situation competently

20. Pessimistic-Optimistic Teacher Behavior

Pessimistic

1. Depressed; unhappy
2. Skeptical
3. Called attention to potential "bad"
4. Expressed hopelessness of "education today," the school system, or fellow educators
5. Noted mistakes; ignored good points
6. Frowned a great deal; had unpleasant facial expression

Optimistic

1. Cheerful; good-natured
2. Genial
3. Joked with pupils on occasion
4. Emphasized potential "good"
5. Looked on bright side; spoke optimistically of the future
6. Called attention to good points; emphasized the positive

21. Immature-Integrated Teacher Behavior

Immature

1. Appeared naive in approach to classroom situations
2. Self-pitying; complaining; demanding
3. Boastful; conceited

Integrated

1. Maintained class as center of activity; kept self out of spotlight; referred to class' activities, not own
2. Emotionally well controlled

22. Narrow-Broad Teacher Behavior

Narrow	Broad
1. Presentation strongly suggested limited background in subject or material; lack of scholarship	1. Presentation suggested good background in subject, good scholarship suggested
2. Did not depart from text	2. Drew examples and explanations from various sources and related fields
3. Failed to enrich discussions with illustrations from related areas	3. Showed evidence of broad cultural background in science, art, literature, history, etc.
4. Showed little evidence of breadth of cultural background in such areas as science, arts, literature, and history	4. Gave satisfying, complete, and accurate answers to questions
5. Answers to pupils' questions incomplete or inaccurate	5. Was constructively critical in approach to subject matter
6. Noncritical approach to subject	

STANFORD TEACHER COMPETENCE APPRAISAL GUIDE

This Appraisal Guide defines the major teacher competences which the program of secondary teacher education at Stanford aims to develop. The total program of teacher education focuses on growth toward these standards as the common target.

To determine whether the program produces the desired growth, levels of competence must be appraised. Evidence for such appraisals may come from the trainee himself, from experienced teachers and administrators who supervise in the schools, from University teachers who instruct the trainees, and from the students taught.

This Appraisal Guide has been designed to assist in a cooperative effort to assess and to improve levels of competence in teaching. The basic sources of evidence are direct observations of the teacher followed by conferences and discussions related to observations. Secondary sources are communications with others who are in a position to observe and to know the teacher's work. The guide encourages the teacher (1) to accept with confidence his proper responsibility for continual self-improvement as a practicing professional in his specialty and (2) to contribute to the ongoing inquiry and the guiding body of theory by which he and his peers seek excellence in their area or specialty subject matter.

Purposely the guide avoids a rigid formula by defining 13 general practitioner competences, around which departmental specialists may build specific standards of expert practices appropriate to subject matter, grade levels, and groupings of students. The teacher being appraised is a most important one of these specialists and should be encouraged to participate in defining and improving standards for his specialty. Self-appraisals followed by observation and conferences with fellow teachers within a department will be useful to teachers as they accept increasing responsibility for self-improvement.

The conference following each observation stresses cooperative sharing of perceptions and ideas between professional teachers focused on the target of improving teaching, supervising, and learning. To facilitate this communication, the conference record is provided in duplicate so both participants may have copies for future use.

STANFORD UNIVERSITY SCHOOL OF EDUCATION

SECONDARY
TEACHER
EDUCATION
PROGRAM

TEACHER COMPETENCE
STANFORD APPRAISAL

STANFORD TEACHER COMPETENCE APPRAISAL GUIDE

AIMS	1	**Clarity of Aims.**	The purposes of the lesson are clear.
	2	**Appropriateness of Aims.**	The aims are neither too easy nor too difficult for the pupils. They are appropriate, and are accepted by the pupils.
PLANNING	3	**Organization of the Lesson.**	The individual parts of the lesson are clearly related to each other in an appropriate way. The total organization facilitates what is to be learned.
	4	**Selection of Content.**	The content is appropriate for the aims of the lesson, the level of the class, and the teaching method.
	5	**Selection of Materials.**	The specific instructional materials and human resources used are clearly related to the content of the lesson and complement the selected method of instruction.
PERFORMANCE	6	**Beginning the Lesson.**	Pupils come quickly to attention. They direct themselves to the tasks to be accomplished.
	7	**Clarity of Presentation.**	The content of the lesson is presented so that it is understandable to the pupils. Different points of view and specific illustrations are used when appropriate.
	8	**Pacing of the Lesson.**	The movement from one part of the lesson to the next is governed by the pupils' achievement. The teacher "stays with the class" and adjusts the tempo accordingly.
	9	**Pupil Participation and Attention.**	The class is attentive. When appropriate the pupils actively participate in the lesson.
	10	**Ending the Lesson.**	The lesson is ended when the pupils have achieved the aims of instruction. There is a deliberate attempt to tie together the planned and chance events of the lesson and relate them to the immediate and long range aims of instruction.
	11	**Teacher-Pupil Rapport.**	The personal relationships between pupils and the teacher are harmonious.
EVALUATION	12	**Variety of Evaluative Procedures.**	The teacher devises and uses an adequate variety of procedures, both formal and informal, to evaluate progress in all of the aims of instruction.
	13	**Use of Evaluation to Improve Teaching and Learning.**	The results of evaluation are carefully reviewed by teacher and pupils for the purpose of improving teaching and learning.
COMMUNITY & PROFESSIONAL	14	**Concern for Professional Standards and Growth.**	The teacher helps, particularly in his specialty, to define and to enforce standards (1) for selecting, training, and licensing of teachers and (2) for working conditions, tools, and equipment necessary for efficient and effective practice.
	15	**Effectiveness in School Staff Relationships.**	The teacher is respectful and considerate of his colleagues. He demonstrates awareness of their personal concerns and professional development.
	16	**Concern for the Total School Program.**	The teacher's concern is not simply for his courses and his students. He sees himself as part of the total school endeavor and actively works with other teachers, students, and administrators to bring about the success of the program.
	17	**Constructive Participation in Community Affairs.**	The teacher understands the particular community context in which he works and helps to translate the purposes of the school's program to the community. He is a responsible member of the community.

OBSERVATION NOTES

Name

Date

Class observed

Length of observation

Observer

0	30%	15%	15%	15%	15%	10%			
0 UNABLE TO OBSERVE	1 WEAK	2 BELOW AVERAGE	3 AVERAGE	4 STRONG	5 SUPERIOR	6 OUTSTANDING	7 TRULY EXCEPTIONAL		
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	1	AIMS
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	2	
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	3	PLANNING
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	4	
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	5	
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	6	PERFORMANCE
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	7	
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	8	
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	9	
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	10	
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	11	
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	12	EVALUATION
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	13	
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	14	COMMUNITY & PROFESSIONAL
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	15	
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	16	
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	17	

GENERAL NOTES

CONFERENCE RECORD

I. Summarize notes relevant to the class setting:

II. Summarize interpretations of observation data: related to aims, planning, performance, and evaluation.

III. Summarize suggestions, possible resources and procedures for improvement of the teaching and learning. What should be retained, discarded, improved? How might this be done? What hypotheses for achieving more effective and efficient teaching and learning were discussed?

Date of Conference ______________ Tentative plans for next observation ______________

TRIAL FORM - SEMANTIC DIFFERENTIAL ATTITUDE SCALE

Name ____

This scale is a new method of measuring what words mean to people. There are no right or wrong answers, so simply indicate your first impressions or feelings about each word or phrase. Try to move fairly rapidly.

At the bottom of the page is a completed example of what you will be asked to do. Please refer to it as you read these instructions.

There are seven blanks separating each word pair. The position of each blank can be described by the following terms. The letter designating it appears above each blank.

Very	Quite	Slightly	Neutral	Slightly	Quite	Very
___	___	___	___	___	___	___

For example, the seven blanks separating the first word pair "Good-Bad" would be read:

	Very Good	Quite Good	Slightly Good	Neutral	Slightly Bad	Quite Bad	Very Bad	
Good	___	___	___	___	___	___	___	Bad

The seven blanks separating the next word pair "Meaningless-Meaningful" would be read:

	Very Meaningless	Quite Meaningless	Slightly Meaningless	Neutral	Slightly Meaningful	Quite Meaningful	Very Meaningful	
aning-less	___	___	___	___	___	___	___	Meaningful

Look at the example below. The word to be described is SISTER. This is not any particular SISTER. Whatever SISTER means to you is what you are to think.

Now let me explain how I described SISTER. Look at the first word pair "Good-Bad." The idea of SISTER was very good to me so I placed an X in the blank closest to "Good." Look at the other word pairs in the example.

Make each item a separate decision. Be sure to put your X on the line provided. Check only one blank for each line. If you make a mistake, erase the X or cross it out and remark.

EXAMPLE

Describe: SISTER

	V	Q	S	N	S	Q	V	
Good	X							Bad
Meaningless							X	Meaningful
Fair		X						Unfair
Negative						X		Positive
Interesting		X						Boring
Unpleasant							X	Pleasant
Necessary							X	Unnecessary
Cruel						X		Kind
Unsuccessful					X			Successful
Valuable	X							Worthless
Reputable	X							Disreputable
Harmful							X	Beneficial

Describe: TESTS

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Negative	___	___	___	___	___	___	___	Positive
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Cruel	___	___	___	___	___	___	___	Kind
Unsuccessful	___	___	___	___	___	___	___	Successful
Valuable	___	___	___	___	___	___	___	Worthless
Reputable	___	___	___	___	___	___	___	Disreputable
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: EDUCATION

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Negative	___	___	___	___	___	___	___	Positive
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Cruel	___	___	___	___	___	___	___	Kind
Unsuccessful	___	___	___	___	___	___	___	Successful
Valuable	___	___	___	___	___	___	___	Worthless
Reputable	___	___	___	___	___	___	___	Disreputable
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: DISCIPLINE

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Negative	___	___	___	___	___	___	___	Positive
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Cruel	___	___	___	___	___	___	___	Kind
Unsuccessful	___	___	___	___	___	___	___	Successful
Valuable	___	___	___	___	___	___	___	Worthless
Reputable	___	___	___	___	___	___	___	Disreputable
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: TEACHING SALARIES

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Negative	___	___	___	___	___	___	___	Positive
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Cruel	___	___	___	___	___	___	___	Kind
Unsuccessful	___	___	___	___	___	___	___	Successful
Valuable	___	___	___	___	___	___	___	Worthless
Reputable	___	___	___	___	___	___	___	Disreputable
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: TEACHER MEETINGS

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Negative	___	___	___	___	___	___	___	Positive
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Cruel	___	___	___	___	___	___	___	Kind
Unsuccessful	___	___	___	___	___	___	___	Successful
Valuable	___	___	___	___	___	___	___	Worthless
Reputable	___	___	___	___	___	___	___	Disreputable
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: PUPILS' PARENTS

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Negative	___	___	___	___	___	___	___	Positive
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Cruel	___	___	___	___	___	___	___	Kind
Unsuccessful	___	___	___	___	___	___	___	Successful
Valuable	___	___	___	___	___	___	___	Worthless
Reputable	___	___	___	___	___	___	___	Disreputable
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: SCHOOL ADMINISTRATORS

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Negative	___	___	___	___	___	___	___	Positive
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Cruel	___	___	___	___	___	___	___	Kind
Unsuccessful	___	___	___	___	___	___	___	Successful
Valuable	___	___	___	___	___	___	___	Worthless
Reputable	___	___	___	___	___	___	___	Disreputable
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: HOMEWORK

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Negative	___	___	___	___	___	___	___	Positive
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Cruel	___	___	___	___	___	___	___	Kind
Unsuccessful	___	___	___	___	___	___	___	Successful
Valuable	___	___	___	___	___	___	___	Worthless
Reputable	___	___	___	___	___	___	___	Disreputable
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: FELLOW TEACHERS

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Negative	___	___	___	___	___	___	___	Positive
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Cruel	___	___	___	___	___	___	___	Kind
Unsuccessful	___	___	___	___	___	___	___	Successful
Valuable	___	___	___	___	___	___	___	Worthless
Reputable	___	___	___	___	___	___	___	Disreputable
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: TEACHING WORK LOAD

	V	Q	S	N	S	Q	V	
Good								Bad
Meaningless								Meaningful
Fair								Unfair
Negative								Positive
Interesting								Boring
Unpleasant								Pleasant
Necessary								Unnecessary
Cruel								Kind
Unsuccessful								Successful
Valuable								Worthless
Reputable								Disreputable
Harmful								Beneficial

Describe: GRADING

	V	Q	S	N	S	Q	V	
Good								Bad
Meaningless								Meaningful
Fair								Unfair
Negative								Positive
Interesting								Boring
Unpleasant								Pleasant
Necessary								Unnecessary
Cruel								Kind
Unsuccessful								Successful
Valuable								Worthless
Reputable								Disreputable
Harmful								Beneficial

Describe: TEACHING AS A CAREER

	V	Q	S	N	S	Q	V	
Good								Bad
Meaningless								Meaningful
Fair								Unfair
Negative								Positive
Interesting								Boring
Unpleasant								Pleasant
Necessary								Unnecessary
Cruel								Kind
Unsuccessful								Successful
Valuable								Worthless
Reputable								Disreputable
Harmful								Beneficial

Describe: STUDENTS

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Negative	___	___	___	___	___	___	___	Positive
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Cruel	___	___	___	___	___	___	___	Kind
Unsuccessful	___	___	___	___	___	___	___	Successful
Valuable	___	___	___	___	___	___	___	Worthless
Reputable	___	___	___	___	___	___	___	Disreputable
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: TEACHER SOCIAL STATUS

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Negative	___	___	___	___	___	___	___	Positive
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Cruel	___	___	___	___	___	___	___	Kind
Unsuccessful	___	___	___	___	___	___	___	Successful
Valuable	___	___	___	___	___	___	___	Worthless
Reputable	___	___	___	___	___	___	___	Disreputable
Harmful	___	___	___	___	___	___	___	Beneficial

PERSONAL DATA FORM--PILOT STUDY

Name____________________ School____________________

Age______ Sex: Male ____________ 01 Female ____________ 02

Single_____ 01 Married_______ 02 Married with children____________ 03

College graduated from:	Washington State University	________ 01
	University of Washington	________ 02
	Central Washington State College	________ 03
	Eastern Washington State College	________ 04
	Western Washington State College	________ 05
	Whitworth	________ 06
	Gonzaga	________ 07
	Whitman	________ 08
	Walla Walla College	________ 09
	Pacific Lutheran College	________ 10
	Puget Sound College	________ 11
	St. Martin's College	________ 12
	Seattle Pacific College	________ 13
	Seattle University	________ 14
	Ft. Wright College	________ 15

Date of graduation (month and year): ____________

Date you began regular teaching (month and year): ____________

College major (30 semester or 45 quarter hours minimum): ____________

College minor(s) (15 semester or 20 quarter hours minimum: ____________

____

Student teaching assignment in -- Major area: ____________ 01

Minor area: ____________ 02

Major and minor areas: ____________ 03

ASSIGNMENT INFORMATION

Teaching grade(s): ____________

Average number of pupils in class(es): ____________

Extra-curricular assignment(s): ____________________________

____

____

Subjects taught: Elementary: All subjects: . . . . ______________ 01

All subjects except art, music, or p.e. . . . ______________ 02

Music . . . . ______________ 03

Art . . . . ______________ 04

P.E. ______________ 05

Foreign Language . . . . ______________ 06

Other - specify . . . . ______________ 07

____

Junior High - specify . . . . ______________

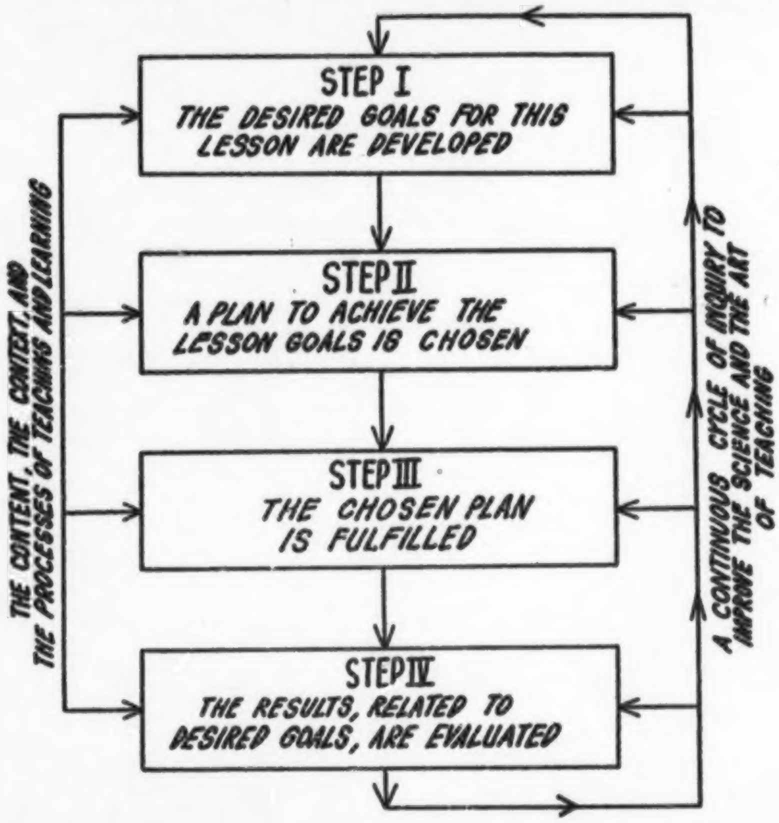
____

____

____

THE SEATTLE TEACHING PERFORMANCE APPRAISAL GUIDE
TO
EXCELLENCE IN CLASSROOM TEACHING
Experimental Form 1965-1966

THE TEACHING ACT - A FOUR STEP CYCLE OF INQUIRY


STEP I
THE DESIRED GOALS FOR THIS LESSON ARE DEVELOPED
STEP II
A PLAN TO ACHIEVE THE LESSON GOALS IS CHOSEN
STEP III
THE CHOSEN PLAN IS FULFILLED
STEP IV
THE RESULTS, RELATED TO DESIRED GOALS, ARE EVALUATED
THE CONTENT, THE CONTEXT, AND THE PROCESSES OF TEACHING AND LEARNING
A CONTINUOUS CYCLE OF INQUIRY TO IMPROVE THE SCIENCE AND THE ART OF TEACHING

The Seattle Teaching Performance Appraisal Guide serves teachers, department heads, principals, and supervisors of beginning teachers, who share responsibilities to improve the teaching abilities it defines. To determine whether improvement in teaching results from any training or professional growth program, performance by the teacher of expected classroom leadership roles in the teaching act, and of the teacher's professional and community participation must be appraised. This guide defines ten expected abilities in the teaching act, and two in professional and community participation. It provides scales to measure observed performance of each ability, supporting evidence in the form of observation notes, and a conference record of improvement suggestions. It adapts to short observation of a specific ability, or to observation of a complete lesson.

The Guide recognizes the importance of the teacher whose performance is appraised, as a professional partner in strengthening the teaching in his field. It does this by expecting a post-observation conference with him and by providing copies for his personal use of the observation notes, the observer's judgment of his performance, and improvement suggestions focused on specific abilities.

Purposely the desired teaching roles are "general," adapting to many different ways to demonstrate them, according to the subject matter, grade level, student grouping, school and community. The emphasis is on the teaching process, not on teacher characteristics. Flexibility and creativity, seeking improved instruction, are built into the Guide design.

The Guide encourages team work among teaching specialists cooperatively to strengthen the practice which is their shared responsibility. Teachers will find appraisals, when made by observers experienced in their field, to be reliable resources in their self-evaluative efforts.

SEATTLE PUBLIC SCHOOLS

PERFORMANCE APPRAISAL GUIDE

STEPS IN THE TEACHING ACT		DESIRED TEACHING ABILITIES—THE GOALS OF PROFESSIONAL GROWTH	
DEVELOPING LESSON GOALS	1	SUITABILITY OF GOALS	The lesson aims are clear—reachable by these students—measurable—show modern knowledge of the subject—relate to what precedes and what follows in the subject—include what and how to learn—serve authorized district educational goals.
	2	STUDENT ACCEPTANCE OF GOALS	The teaching explores student understanding of lesson aims—relates aims meaningfully to the present lives of these students—the students see the goals as worth working to achieve—the teaching helps students to establish personal goals consistent with lesson aims.
PLANNING GOAL ACTION	3	EXPLORATION OF HUMAN AND MATERIAL RESOURCES	The teaching appraises student talents, activities, interests—stresses the students as primary resources, fellow teachers and staff specialists, parents, other significant adults—available space, texts, tools, audio-visual aids, labs, libraries—time, staff, and material budget limits.
	4	SELECTING THE PLAN FOR THIS CLASS	The teaching considers alternative ways to use available resources to accomplish aims—selects a plan promising optimal success within budget limits of available resources and conforming to school policies—the teaching helps the students to see how the plan makes sense—to plan their own learning activities—inquiry, not habit, guides the plan choice.
	5	ORGANIZING THE CLASS TO ACHIEVE THE PLAN	The teaching clearly defines who is to do what-when-why-how as the plan unfolds—each student is an active, valued member of the organization—each student is helped to see how he fits, how he belongs, how he can be useful, what to expect of the teacher, of himself, and of other students—ground rules are established.
FULFILLING THE PLAN	6	CLASSROOM CONTROL EFFECTIVE ACTION	The teaching follows the plan—each phase has an introduction, a body of action, and a conclusion—unnecessary deviations are controlled—clear, intriguing, strategically timed explanations, demonstrations, reminders gain attention—maintain interest—motivate—inform of plan progress—encourage student initiative and self-discipline—materials are ready when needed—ground rules are enforced.
	7	CLASSROOM CLIMATE EFFICIENT ACTION	The teaching conserves human and material resources-people-property-time—shows sensitivity to and understanding of attention span-fatigue-human problems—pacing adapts to student achievement—self-respect and confidence is encouraged—rapport is positive—problems, not people, are attacked—leadership patterns are democratic.
	8	ACTIVE STUDENT PARTICIPATION	The teaching delegates to students responsibilities they can handle—encourages students to teach themselves, to help teach others—leaves something for the students to doubt, to ask, to investigate, to test, to interpret, to express, to discover, to be responsible for, to recognize as theirs.
EVALUATING RESULTS	9	MEASURING GOAL ACHIEVEMENT AND COSTS	The teaching measures what was achieved and "how we did it"—using modern techniques for this subject—appraises costs to the participants—helps students to design their own tests—to investigate their own progress—measurements are timed to serve the next step in the class effort.
	10	USING MEASUREMENTS TO IMPROVE TEACHING AND LEARNING	The teaching uses test scores, grades to guide the teaching and learning decisions—not as ends in themselves—post-test discussions are learning experiences, for both students and the teachers—measurements are not used to label students, to indoctrinate fear of failure, to develop uncritical worship of high grades.
PROFESSIONAL AND COMMUNITY	11	PROFESSIONAL PARTICIPATION	The teacher accepts with his fellows responsibility to define and enforce standards admitting beginners to practice in his field—to achieve the in-service conditions, training opportunities, and rewards which are essential to the improvement of practice; to a professional career commitment.
	12	COMMUNITY PARTICIPATION	The teacher studies continuously his school community—relating his professional services to its educational problems. He helps to clarify and strengthen the educational values and expectations of parents and community leadership, related to his special field.

OBSERVED PERFORMANCE LEVEL

LOW ←——RANGE OF SEATTLE PRACTICE——→ HIGH

PERFORMANCE IN UPPER 25%
IN MIDDLE RANGE
IN LOWER 30%
NOT OBSERVED

	30%	NEXT 15%	NEXT 15%	NEXT 15%	NEXT 15%	NEXT 5%	NEXT 5%		
0	1	2	3	4	5	6	7		
☐	☐	☐	☐	☐	☐	☐	☐	1	DEVELOPING LESSON GOALS
☐	☐	☐	☐	☐	☐	☐	☐	2	
☐	☐	☐	☐	☐	☐	☐	☐	3	PLANNING GOAL ACTION
☐	☐	☐	☐	☐	☐	☐	☐	4	
☐	☐	☐	☐	☐	☐	☐	☐	5	
☐	☐	☐	☐	☐	☐	☐	☐	6	FULFILLING THE PLAN
☐	☐	☐	☐	☐	☐	☐	☐	7	
☐	☐	☐	☐	☐	☐	☐	☐	8	
☐	☐	☐	☐	☐	☐	☐	☐	9	EVALUATING RESULTS
☐	☐	☐	☐	☐	☐	☐	☐	10	
☐	☐	☐	☐	☐	☐	☐	☐	11	PROFESSIONAL AND COMMUNITY
☐	☐	☐	☐	☐	☐	☐	☐	12	

OBSERVATION NOTES

Teacher______________________________

School____________________ **Class**__________

Length of observation__________ **Date**__________

Observer______________________________

SEATTLE PUBLIC SCHOOLS

PERFORMANCE APPRAISAL CONFERENCE RECORD

Date________________________ Participants____

Descriptive information about this teaching situation:

Previous observations, conferences, if any, supporting this performance appraisal:

Dates of observation	Followed by conference?	Copy furnished teacher?

Suggestions for action to improve teaching, (by the teacher, by others, by the teacher to help others)

1. **To strengthen development of goals:**

2. **To strengthen planning:**

3. **To strengthen plan fulfillment:**

4. **To strengthen evaluation:**

5. **Professional and community:**

______________________________ ______________________________
Teacher's Signature **Observer's Signature**

BE SURE A COPY GOES TO THE TEACHER

SEMANTIC DIFFERENTIAL ATTITUDE SCALE

This scale is a new method of measuring what words mean to people. There are no right or wrong answers, so simply indicate your first impressions or feelings about each word or phrase. Try to move fairly rapidly.

At the bottom of the page is a completed example of what you will be asked to do. Please refer to it as you read these instructions.

There are seven blanks separating each word pair. The position of each blank can be described by the following terms. The letter designating it appears above each blank.

Very	Quite	Slightly	Neutral	Slightly	Quite	Very
____	____	____	____	____	____	____

For example, the seven blanks separating the first word pair "Good-Bad" would be read:

	Very Good	Quite Good	Slightly Good	Neutral	Slightly Bad	Quite Bad	Very Bad	
Good	____	____	____	____	____	____	____	Bad

The seven blanks separating the next word pair "Meaningless-Meaningful: would be read:

	Very Meaning-less	Quite Meaning-less	Slightly Meaning-less	Neutral	Slightly Meaning-ful	Quite Meaning-ful	Very Meaning-ful	
Meaning-less	____	____	____	____	____	____	____	Meaning-ful

Look at the example below. The word to be described is TAXPAYER. This is not any particular TAXPAYER. Whatever TAXPAYER means to you is what you are to think.

Now let me explain how I described TAXPAYER. Look at the first word pair "Good-Bad." The idea of TAXPAYER was very good to me so I placed an X in the blank closest to "Good." Look at the other word pairs in the example.

Make each item a separate decision. Be sure to put your X on the line provided. Check only one blank for each line. If you make a mistake, erase the X or cross it out and remark.

EXAMPLE

Describe: TAXPAYER

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Valuable	___	___	___	___	___	___	___	Worthless
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: TESTS

	V	Q	S	N	S	Q	V	
Good								Bad
Meaningless								Meaningful
Fair								Unfair
Interesting								Boring
Unpleasant								Pleasant
Necessary								Unnecessary
Valuable								Worthless
Harmful								Beneficial

Describe: TEACHER MEETINGS

	V	Q	S	N	S	Q	V	
Good								Bad
Meaningless								Meaningful
Interesting								Boring
Unpleasant								Pleasant
Necessary								Unnecessary
Unsuccessful								Successful
Valuable								Worthless
Harmful								Beneficial

Describe: SCHOOL ADMINISTRATORS

	V	Q	S	N	S	Q	V	
Good								Bad
Meaningless								Meaningful
Fair								Unfair
Negative								Positive
Unsuccessful								Successful
Valuable								Worthless
Harmful								Beneficial

Describe: HOMEWORK

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Negative	___	___	___	___	___	___	___	Positive
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Valuable	___	___	___	___	___	___	___	Worthless
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: TEACHING WORK LOAD

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Fair	___	___	___	___	___	___	___	Unfair
Interesting	___	___	___	___	___	___	___	Boring
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Valuable	___	___	___	___	___	___	___	Worthless
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: GRADING

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Unpleasant	___	___	___	___	___	___	___	Pleasant
Necessary	___	___	___	___	___	___	___	Unnecessary
Valuable	___	___	___	___	___	___	___	Worthless
Harmful	___	___	___	___	___	___	___	Beneficial

Describe: TEACHING AS A CAREER

	V	Q	S	N	S	Q	V	
Good	___	___	___	___	___	___	___	Bad
Meaningless	___	___	___	___	___	___	___	Meaningful
Fair	___	___	___	___	___	___	___	Unfair
Unpleasant	___	___	___	___	___	___	___	Pleasant
Unsuccessful	___	___	___	___	___	___	___	Successful
Valuable	___	___	___	___	___	___	___	Worthless

APPENDIX B

SUMMARY TABLES

TABLE 25

DATA FROM THE TRIAL FORM OF THE SEMANTIC DIFFERENTIAL ATTITUDE SCALE

1. Concept: Tests

Adjective Pairs	Round 1		Round 2		Round 3	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Good-Bad	2.52	1.24	2.59	1.40	2.52	1.54
Meaningful-Meaningless	2.67	1.47	2.77	1.41	2.86	1.46
Fair-Unfair	2.85	1.32	2.82	1.44	3.05	1.57
Positive-Negative	3.12	1.42	2.90	1.41	3.14	1.49
Interesting-Boring	2.44	1.26	3.00	1.48	3.05	1.39
Pleasant-Unpleasant	3.78	1.42	4.05	1.56	3.76	1.67
Necessary-Unnecessary	2.00	1.64	2.33	1.39	2.00	1.52
Kind-Cruel	3.74	1.26	3.86	1.13	3.57	1.36
Successful-Unsuccessful	2.85	1.06	2.95	1.50	3.05	1.53
Valuable-Worthless	2.33	1.41	2.59	1.50	2.62	1.60
Reputable-Disreputable	3.07	1.36	3.18	1.37	2.95	1.53
Beneficial-Harmful	2.15	.82	2.64	1.18	2.81	1.40

2. Concept: Education

Adjective Pairs	Round 1 Mean	Round 1 S.D.	Round 2 Mean	Round 2 S.D.	Round 3 Mean	Round 3 S.D.
Good-Bad	1.07	.27	1.27	.46	1.14	.36
Meaningful-Meaningless	1.15	.46	1.32	.57	1.14	.36
Fair-Unfair	1.93	1.27	2.23	1.27	2.10	.94
Positive-Negative	1.74	.98	1.73	.99	1.62	.67
Interesting-Boring	1.37	.57	1.45	.67	1.57	.51
Pleasant-Unpleasant	1.63	.63	1.68	.72	1.86	.79
Necessary-Unnecessary	2.00	1.64	2.33	1.40	2.00	1.52
King-Cruel	2.63	1.15	2.77	1.24	2.81	1.25
Successful-Unsuccessful	1.81	.74	1.82	.91	1.90	1.00
Valuable-Worthless	1.15	.46	1.14	.35	1.05	.22
Reputable-Disreputable	1.78	1.01	1.82	1.18	1.67	.97
Beneficial-Harmful	1.11	.32	1.23	.53	1.14	.36

TABLE 25 - Continued

3. Concept: Discipline

Adjective Pairs	Round 1		Round 2		Round 3	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Good-Bad	1.63	.84	1.77	1.11	1.95	1.32
Meaningful-Meaningless	1.59	.69	1.82	.67	2.00	1.14
Fair-Unfair	2.11	1.05	2.05	1.02	2.38	.87
Positive-Negative	2.30	1.33	2.50	1.74	2.65	1.66
Interesting-Boring	3.33	1.04	3.68	1.17	3.48	1.12
Pleasant-Unpleasant	3.96	1.87	4.45	1.57	4.24	1.92
Necessary-Unnecessary	1.32	.56	1.32	.48	1.62	.92
Kind-Cruel	2.85	1.17	3.41	1.26	3.29	1.23
Successful-Unsuccessful	2.44	1.12	2.45	1.22	2.57	1.33
Valuable-Worthless	1.70	.91	1.82	1.10	1.81	1.08
Reputable-Disreputable	2.74	1.43	2.91	1.31	2.95	1.16
Beneficial-Harmful	1.74	1.32	1.64	1.05	2.05	1.12

4. Concept: Teaching Salaries

Good-Bad	2.78	1.37	2.82	1.44	2.90	1.34
Meaningful-Meaningless	2.37	1.50	2.32	1.39	2.52	1.33
Fair-Unfair	2.85	1.52	3.18	1.40	3.19	1.57
Positive-Negative	2.85	1.26	2.90	1.45	2.81	.98
Interesting-Boring	3.00	1.21	3.09	1.07	2.90	1.09
Pleasant-Unpleasant	2.56	1.37	2.82	1.44	2.71	1.19
Necessary-Unnecessary	1.26	.71	1.27	.88	1.19	.40
Kind-Cruel	3.00	1.30	3.09	1.31	3.10	1.09
Successful-Unsuccessful	2.67	1.33	2.95	1.40	3.14	1.15
Valuable-Worthless	1.56	.93	1.50	.80	1.67	1.02
Reputable-Disreputable	2.56	1.34	2.82	1.37	3.14	1.01
Beneficial-Harmful	1.74	1.32	1.64	1.05	2.05	1.12

TABLE 25 - Continued

5. Concept: Teacher Meetings

Adjective Pairs	Round 1		Round 2		Round 3	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Good-Bad	3.04	1.58	3.00	1.54	2.86	1.28
Meaningful-Meaningless	2.96	1.70	3.09	1.60	2.81	1.50
Fair-Unfair	3.04	1.29	3.27	1.24	2.81	1.08
Positive-Negative	2.81	1.17	3.23	1.54	2.90	1.18
Interesting-Boring	3.31	1.49	3.45	1.90	3.10	1.48
Pleasant-Unpleasant	2.88	1.34	2.91	1.48	2.90	1.45
Necessary-Unnecessary	1.74	1.02	2.00	1.02	2.05	.87
Kind-Cruel	3.44	1.09	3.36	1.14	3.43	.98
Successful-Unsuccessful	2.70	1.44	3.09	1.63	2.67	1.11
Valuable-Worthless	2.22	1.16	2.91	1.48	2.33	1.16
Reputable-Disreputable	2.74	1.29	3.00	1.38	3.19	1.08
Beneficial-Harmful	2.33	1.27	2.50	1.06	2.52	1.03
6. Concept: Pupils' Parents						
Good-Bad	2.56	1.28	2.27	1.16	2.38	1.16
Meaningful-Meaningless	2.30	1.77	2.36	1.40	2.05	.87
Fair-Unfair	2.74	1.20	2.91	1.45	2.86	1.28
Positive-Negative	2.81	1.47	2.73	1.58	3.10	1.61
Interesting-Boring	2.00	1.11	1.95	1.09	1.90	.89
Pleasant-Unpleasant	2.41	1.37	2.45	.86	2.29	1.06
Necessary-Unnecessary	1.30	.72	1.52	1.08	1.33	.73
Kind-Cruel	2.78	1.31	2.43	1.20	2.90	1.38
Successful-Unsuccessful	2.85	1.32	3.23	1.31	3.29	1.35
Valuable-Worthless	1.59	.97	1.91	1.27	1.95	1.07
Reputable-Disreputable	2.81	1.33	2.82	1.18	2.67	1.20
Beneficial-Harmful	2.48	1.48	2.73	1.32	2.52	1.29

TABLE 25 - Continued

7. Concept: School Administrators

Adjective Pairs	Round 1		Round 2		Round 3	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Good-Bad	1.93	1.04	2.14	1.13	2.10	.94
Meaningful-Meaningless	1.96	.98	2.14	1.25	2.24	1.18
Fair-Unfair	2.11	1.09	2.23	1.19	2.14	1.01
Positive-Negative	2.44	1.25	2.41	1.26	2.43	1.20
Interesting-Boring	2.37	1.33	2.41	1.05	2.62	1.40
Pleasant-Unpleasant	2.00	.96	2.24	1.18	2.10	1.02
Necessary-Unnecessary	1.11	.32	1.41	.90	1.38	.59
Kind-Cruel	2.11	1.09	2.55	1.30	2.38	1.07
Successful-Unsuccessful	2.04	1.09	2.23	1.15	2.25	1.02
Valuable-Worthless	1.70	.72	1.76	1.00	1.75	.79
Reputable-Disreputable	2.04	1.29	2.23	1.34	1.85	1.18
Beneficial-Harmful	1.70	.99	1.91	1.11	2.05	1.05

8. Concept: Home Work

Adjective Pairs	Round 1		Round 2		Round 3	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Good-Bad	2.33	1.39	2.55	1.47	2.52	1.37
Meaningful-Meaningless	2.19	.92	2.36	1.29	2.81	1.47
Fair-Unfair	2.93	1.41	2.64	1.22	2.67	1.16
Positive-Negative	2.48	1.25	2.55	1.14	2.81	1.12
Interesting-Boring	2.81	1.30	2.64	1.33	2.86	1.01
Pleasant-Unpleasant	3.52	1.40	3.27	1.28	3.29	1.31
Necessary-Unnecessary	2.35	1.52	2.52	1.47	2.52	1.63
Kind-Cruel	3.37	1.08	3.48	.98	3.52	.93
Successful-Unsuccessful	2.85	1.46	2.95	1.16	2.81	1.17
Valuable-Worthless	2.41	1.39	2.43	1.29	2.43	1.36
Reputable-Disreputable	3.26	1.48	3.10	1.26	3.29	1.31
Beneficial-Harmful	2.37	1.31	2.19	1.08	2.43	1.08

TABLE 25 - Continued

9. Concept: Fellow Teachers

Adjective Pairs	Round 1		Round 2		Round 3	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Good-Bad	1.56	.97	1.64	1.05	1.81	1.03
Meaningful-Meaningless	1.96	1.09	1.95	1.25	2.00	1.14
Fair-Unfair	2.15	1.29	2.00	1.07	2.29	1.42
Positive-Negative	2.11	1.34	1.95	1.17	2.14	1.01
Interesting-Boring	1.85	1.06	1.86	1.36	1.81	1.12
Pleasant-Unpleasant	1.63	.63	1.64	.95	1.71	1.01
Necessary-Unnecessary	1.48	.85	1.32	.78	1.57	.93
Kind-Cruel	2.12	.99	1.77	.92	1.90	1.04
Successful-Unsuccessful	2.11	.89	2.05	1.21	2.19	1.08
Valuable-Worthless	1.89	1.22	1.64	1.00	1.76	.94
Reputable-Disreputable	2.22	1.31	1.82	1.18	1.81	1.08
Beneficial-Harmful	1.70	.82	1.82	.90	1.86	1.06

10. Concept: Teaching Work Load

Good-Bad	3.19	1.47	3.55	1.87	3.48	1.81
Meaningful-Meaningless	2.78	1.22	2.59	1.05	2.76	1.00
Fair-Unfair	3.00	1.49	3.36	1.47	3.33	1.71
Positive-Negative	3.00	1.27	3.08	1.22	3.38	1.43
Interesting-Boring	2.35	1.13	2.55	1.14	2.71	1.27
Pleasant-Unpleasant	2.81	1.30	2.82	1.37	3.62	1.72
Necessary-Unnecessary	1.81	.92	2.36	1.53	2.57	1.54
Kind-Cruel	3.00	1.24	3.55	1.26	3.62	1.24
Successful-Unsuccessful	2.70	1.17	3.41	1.33	3.05	1.20
Valuable-Worthless	2.30	1.20	2.95	1.17	2.95	1.16
Reputable-Disreputable	3.11	1.40	3.27	1.12	3.29	1.19
Beneficial-Harmful	2.74	1.48	3.05	1.43	3.19	1.75

TABLE 25 - Continued

11. Concept: Grading

Adjective Pairs	Round 1		Round 2		Round 3	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Good-Bad	3.44	1.63	3.41	1.94	3.19	1.60
Meaningful-Meaningless	3.11	1.60	3.32	1.78	3.10	1.61
Fair-Unfair	3.58	1.60	3.73	1.55	3.29	1.38
Positive-Negative	3.78	1.63	3.59	1.53	3.38	1.77
Interesting-Boring	3.44	1.42	3.41	1.44	3.38	1.28
Pleasant-Unpleasant	4.40	1.47	4.52	1.72	3.95	1.53
Necessary-Unnecessary	2.44	1.48	2.59	1.74	2.43	1.43
Kind-Cruel	2.63	1.31	2.50	1.30	2.71	1.19
Successful-Unsuccessful	3.56	1.45	3.86	1.42	3.29	1.49
Valuable-Worthless	2.85	1.51	3.41	1.65	2.81	1.40
Reputable-Disreputable	3.41	1.19	3.64	1.26	3.29	1.15
Beneficial-Harmful	3.44	1.63	3.59	1.71	3.19	1.60

12. Concepts: Teaching as a Career

Adjective Pairs	Round 1		Round 2		Round 3	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Good-Bad	1.56	.89	1.68	.78	1.52	.68
Meaningful-Meaningless	1.52	.70	1.64	.85	1.67	.91
Fair-Unfair	2.11	1.12	2.18	1.14	2.19	1.08
Positive-Negative	1.89	1.09	1.95	1.05	1.95	1.07
Interesting-Boring	1.37	.63	1.45	.86	1.62	1.16
Pleasant-Unpleasant	1.63	1.08	1.77	.87	1.67	.58
Necessary-Unnecessary	1.85	1.20	1.77	1.27	1.86	1.39
Kind-Cruel	2.63	1.31	2.50	1.30	2.71	1.19
Successful-Unsuccessful	1.93	1.07	2.14	1.08	1.86	.91
Valuable-Worthless	1.33	.68	1.45	.96	1.67	1.07
Reputable-Disreputable	1.78	1.34	1.86	1.46	1.76	1.14
Beneficial-Harmful	1.78	1.37	1.64	.85	1.90	1.04

TABLE 25 - Continued

13. Concepts: Students

Adjective Pairs	Round 1		Round 2		Round 3	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Good-Bad	2.26	1.06	2.00	1.20	1.62	.59
Meaningful-Meaningless	1.93	1.00	1.86	1.17	1.81	1.03
Fair-Unfair	2.48	1.22	2.55	1.22	2.14	1.01
Positive-Negative	2.56	1.34	2.27	1.24	2.14	1.01
Interesting-Boring	1.38	.57	1.27	.70	1.33	.48
Pleasant-Unpleasant	1.96	1.16	2.27	1.32	1.81	.75
Necessary-Unnecessary	1.74	1.16	1.64	1.22	1.35	.93
Kind-Cruel	2.37	1.21	2.41	1.22	2.62	1.31
Successful-Unsuccessful	2.33	1.00	2.64	1.40	2.47	1.03
Valuable-Worthless	1.70	.99	1.95	1.33	1.90	1.04
Reputable-Disreputable	2.56	1.34	2.55	1.26	2.43	1.17
Beneficial-Harmful	2.15	1.20	2.45	1.30	2.29	1.27

14. Concepts: Teachers' Social Status

Good-Bad	2.19	.96	2.18	.85	2.14	1.11
Meaningful-Meaningless	2.59	1.56	2.73	1.49	2.38	1.36
Fair-Unfair	3.19	1.57	2.82	1.30	2.43	1.33
Positive-Negative	2.96	1.45	2.73	1.20	2.57	1.40
Interesting-Boring	2.81	1.21	2.86	.94	2.52	1.21
Pleasant-Unpleasant	2.69	1.46	3.00	1.07	2.76	1.34
Necessary-Unnecessary	2.81	1.39	3.00	1.51	2.90	1.70
Kind-Cruel	3.15	1.23	3.00	1.23	3.57	1.17
Successful-Unsuccessful	2.96	1.29	2.73	1.12	2.52	1.17
Valuable-Worthless	2.89	1.31	2.91	1.41	2.62	1.43
Reputable-Disreputable	2.56	1.58	2.36	1.29	2.29	1.45
Beneficial-Harmful	2.56	1.28	2.77	1.07	2.57	1.60

TABLE 26

SUMMARY OF DATA ON PILOT STUDY TEACHER RATINGS AS RELATED TO SUBJECTS' CHARACTERISTICS

Code #	Mn Score Stanford	Rank	Mn Score Ryans	Rank	Grade Level	Student Teacher Grade	Cum GPA	Ed GPA	Major
1	2.02	24	3.46	22	6	B	2.5	3.2	Soc. St.
2	2.87	14.5	4.69	6	Int-Adj	A	2.8	2.8	Soc. St.
3	3.38	3	3.60	20	3	A	3.1	3.6	Soc. St.
4	2.29	23	3.20	25	3	A	2.8	3.2	Soc. St.
5	3.16	7	4.12	14	1	A	2.4	3.0	Soc. St.
6	2.88	13	4.91	2	5	A	2.8	3.2	Music
7	2.67	18	4.58	7	3	A	2.9	3.4	Soc. St.
8	2.56	19.5	3.40	23	1	B	2.6	3.3	Lang Arts
9	2.56	19.5	3.87	16	4	B	2.9	3.3	Soc. St.
10	2.34	22	3.17	26	8-9	B	2.7	2.5	Art
11	Dropped	from study							
12	3.32	4	4.20	12	7-9	B	2.8	2.9	Soc. St.
13	Dropped	from Study							
14	2.91	12	4.81	4	1	B	2.6	3.3	Lang Arts
15	2.53	21	3.76	17	1	A	3.0	3.3	Soc. St.
16	4.08	2	5.83	1	1	A	2.8	3.4	Lang Arts
17	3.03	9	4.21	10	K	B	3.1	3.0	Soc. St.
18	2.87	14.5	3.33	24	K	A	2.3	2.2	Lang Arts
19	3.23	5	4.40	9	6	B	2.1	2.4	
20	1.82	26	3.91	15	4	B	3.3	3.2	Soc. St.
21	3.05	8	4.16	13	5	A	3.4	3.7	Soc. St.
22	2.79	17	3.74	18	3	B	3.0	3.2	Soc. St.

TABLE 26 - Continued

Code	Mn Score Stanford	Rank	Mn Score Ryans	Rank	Grade Level	Student Teacher Grade	Cum GPA	Ed GPA	Major
23	3.00	10	3.59	21	8	B	3.2	3.3	Music
24	2.92	11	4.57	8	3	A	2.5	2.7	H. Ec.
25	3.19	6	4.26	11	2	A	3.7	3.8	Lang Arts
26	4.49	1	4.78	5	Int.	A	2.9	3.6	Soc. St.
27	2.81	16	4.87	3	3	B	3.3	3.6	Lang Arts
28	1.97	25	3.73	19	11	A	2.2	3.0	Chem.

74.74
Mn= 2.87
N = 26

107.15
Mn = 4.12
N = 26

TABLE 27
CORRELATIONS BETWEEN OBSERVER PAIRS ON ALL ROUNDS

Observer Pairs	Scale Item										Total
	1	2	3	4	5	6	7	8	9	10	
					Round 1						
AB	.495	.485	.466	.385	.335	.704	.451	.545	.346	.217	.459
AC	.257	.205	.284	.415	.258	.200	.255	.394	.399	.184	.290
BC	.452	.601	.731	.701	.415	.551	.669	.645	.678	.461	.589
DE	.671	.773	.713	.743	.754	.786	.807	.800	.621	.658	.720
DF	.434	.421	.580	.601	.538	.575	.575	.383	.293	.363	.475
EF	.441	.353	.684	.525	.500	.554	.476	.383	.354	.342	.455
GH	.664	.746	.644	.652	.665	.668	.711	.706	.792	.775	.692
GI	.610	.586	.533	.607	.480	.512	.555	.427	.573	.443	.524
HI	.613	.664	.545	.632	.450	.562	.544	.469	.508	.389	.537

TABLE 27 - Continued

Observer Pairs	Scale Item										Total
	1	2	3	4	5	6	7	8	9	10	
					Round 2						
AD	.420	.601	.526	.402	.593	.407	.443	.452	.710	.646	.517
AG	.575	.765	.692	.501	.724	.701	.619	.708	.654	.693	.657
DG	.651	.827	.773	.790	.845	.803	.823	.669	.785	.795	.771
BE	.765	.761	.724	.783	.798	.722	.720	.750	.603	.544	.723
BH	.689	.670	.675	.722	.815	.815	.857	.705	.607	.498	.719
EH	.594	.780	.815	.729	.726	.725	.734	.649	.615	.632	.706
CF	.278	.411	.372	.324	.503	.572	.511	.363	.319	.334	.416
CI	.115	.010	.129	.088	.208	.044	.194	.164	.268	.061	.112
FI	.317	.247	.321	.166	.233	.298	.159	.346	.035	.342	.260

TABLE 27 - Continued

Observer Pairs	Scale Item										Total
	1	2	3	4	5	6	7	8	9	10	
					Round 3						
AB	.834	.901	.752	.845	.865	.877	.709	.832	.866	.767	.819
AH	.776	.888	.784	.795	.803	.876	.794	.839	.864	.765	.815
BH	.794	.780	.803	.764	.879	.877	.885	.766	.817	.872	.825
CE	.224	.072	.179	.222	.394	.450	.488	.170	.338	.264	.292
CF	.118	.168	.394	.200	.514	.391	.433	.181	.285	.284	.301
EF	.467	.533	.508	.598	.662	.686	.637	.488	.249	.411	.541
DG	.760	.666	.688	.728	.679	.752	.680	.738	.608	.653	.691
DI	.607	.627	.454	.512	.582	.493	.603	.453	.445	.520	.507
GI	.496	.470	.452	.462	.537	.441	.696	.561	.543	.558	.505

TABLE 27 - Continued

Observer Pairs	Scale Item 1	2	3	4	5	6	7	8	9	10	Total
					Round 4						
AC	.490	.744	.509	.574	.653	.799	.783	.596	.668	.668	.648
AF	.158	.253	.304	.245	.542	.441	.152	.175	.000	.132	.266
CF	.304	.390	.512	.259	.678	.509	.368	.379	.415	.525	.435
BH	.682	.754	.657	.672	.732	.737	.702	.758	.683	.669	.706
BI	.804	.864	.840	.809	.805	.778	.761	.790	.887	.864	.815
HI	.780	.731	.769	.819	.695	.791	.796	.692	.766	.778	.757
DE	.797	.822	.804	.819	.828	.822	.799	.707	.785	.780	.786
DG	.823	.829	.869	.778	.869	.891	.856	.868	.890	.779	.843
EG	.801	.785	.899	.828	.878	.868	.843	.818	.746	.834	.822

TABLE 28

SUMMARY OF DATA ON THE SEMANTIC DIFFERENTIAL ATTITUDE SCALE BY TREATMENT GROUPS AND ADMINISTRATIONS (N = 96)

Treatment Groups	Round 1		Round 2		Round 3		Round 4	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
A. Concept: "Tests"								
1. Adjective pair: "good-bad"								
Group 1	2.27	1.15	2.56	1.16	2.56	1.04	2.35	.65
Group 2	2.64	1.15	2.44	.93	2.70	.78	2.41	1.12
Group 3	2.44	1.26	2.32	1.09	2.52	1.12	2.60	1.04
Group 4	2.15	1.08	2.31	.93	2.36	.64	2.40	.71
2. Adjective pair: "meaningful-meaningless"								
Group 1	2.15	1.08	2.63	1.12	2.76	1.30	2.50	1.02
Group 2	2.92	1.47	2.96	1.43	3.07	1.33	2.46	1.21
Group 3	2.52	1.23	2.57	1.12	2.36	.76	2.88	1.39
Group 4	2.19	1.13	3.27	1.89	2.56	1.08	2.79	1.50
3. Adjective pair: "fair-unfair"								
Group 1	2.50	1.03	3.04	1.22	2.72	.94	2.48	.79
Group 2	2.84	1.11	3.04	1.37	2.96	1.13	2.88	1.07
Group 3	2.68	1.25	2.95	1.53	2.72	1.43	3.08	1.29
Group 4	2.88	1.28	3.08	1.35	2.79	1.14	2.70	.93
4. Adjective pair: "interesting-boring"								
Group 1	2.36	1.22	2.52	1.16	2.44	1.16	2.54	1.38
Group 2	2.54	1.10	2.81	1.00	2.96	1.22	2.81	1.13
Group 3	2.48	1.30	2.57	.99	2.72	1.17	2.64	1.08
Group 4	2.73	1.31	2.73	1.28	3.12	1.36	2.83	1.11
5. Adjective pair: "pleasant-unpleasant"								
Group 1	3.71	1.22	3.46	1.14	3.44	1.39	3.42	1.10
Group 2	4.24	1.30	4.22	1.25	4.15	1.17	3.54	1.30
Group 3	4.12	1.17	4.32	1.17	3.63	1.35	3.72	1.14
Group 4	4.08	1.62	3.81	1.17	3.92	1.38	4.22	1.28
6. Adjective pair: "necessary-unnecessary"								
Group 1	1.81	.80	1.81	.92	1.80	1.00	2.09	1.31
Group 2	1.80	.82	2.00	.67	2.26	1.02	2.50	1.68
Group 3	1.72	1.10	2.17	1.34	2.20	1.23	2.12	1.05
Group 4	1.96	1.11	2.38	1.55	2.24	.93	2.35	1.15

TABLE 28--Continued

Treatment Groups	Round 1		Round 2		Round 3		Round 4	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
7. Adjective pair: "valuable-worthless"								
Group 1	2.15	.79	2.37	.97	2.32	.95	2.48	.95
Group 2	2.28	.74	2.32	.77	2.74	.90	2.23	.86
Group 3	2.28	.89	2.13	.82	2.16	.69	2.26	.79
Group 4	2.15	1.01	2.23	.91	2.40	1.00	2.61	1.03
8. Adjective pair: "beneficial-harmful"								
Group 1	2.35	1.09	2.52	1.12	2.52	1.09	2.58	.83
Group 2	2.56	.92	2.74	1.02	3.00	1.00	2.77	.99
Group 3	2.28	1.06	2.74	1.32	2.96	1.37	2.60	1.12
Group 4	2.31	1.12	2.54	1.14	2.68	1.07	2.78	1.00
B. Concept: "Teacher Meetings"								
1. Adjective pair: "good-bad"								
Group 1	2.04	1.29	2.19	.92	2.08	1.15	2.63	1.25
Group 2	2.27	.88	2.54	.96	2.56	.97	2.52	1.05
Group 3	2.44	1.26	2.57	1.27	2.52	1.45	2.76	1.45
Group 4	2.12	1.40	2.15	.93	2.38	1.10	2.12	.60
2. Adjective pair: "meaningful-meaningless"								
Group 1	2.26	1.32	2.33	1.21	2.72	1.75	2.54	1.29
Group 2	2.42	.86	2.61	1.03	2.78	1.45	2.31	1.01
Group 3	2.28	1.02	2.78	1.57	2.68	1.65	2.72	1.46
Group 4	2.31	1.46	2.54	1.56	2.32	.85	2.67	1.34
3. Adjective pair: "interesting-boring"								
Group 1	2.59	1.22	2.67	1.24	2.52	1.39	2.58	1.32
Group 2	3.19	1.30	3.32	1.31	3.15	1.26	2.88	1.28
Group 3	2.92	1.61	2.91	1.73	3.08	1.55	3.12	1.64
Group 4	2.73	1.49	2.96	1.56	2.80	1.29	2.79	1.44
4. Adjective pair: "pleasant-unpleasant"								
Group 1	2.37	1.08	2.56	1.09	2.56	1.29	2.38	1.10
Group 2	2.73	1.22	2.57	.88	2.74	.86	2.65	1.33
Group 3	2.88	1.33	3.26	1.63	3.16	1.43	3.00	1.29
Group 4	2.38	1.06	3.04	1.37	2.80	1.19	2.92	1.21
5. Adjective pair: "necessary-unnecessary"								
Group 1	1.59	1.08	1.93	1.07	1.84	1.03	2.00	1.29
Group 2	2.12	.99	2.25	1.18	1.93	.73	1.85	.68
Group 3	2.04	1.46	2.09	1.28	2.16	1.43	2.28	1.43
Group 4	1.73	1.12	2.04	1.34	2.24	1.13	2.25	1.26

TABLE 28 --Continued

Treatment Groups	Round 1		Round 2		Round 3		Round 4	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
6. Adjective pair: "valuable-worthless"								
Group 1	2.11	1.19	2.15	.91	2.16	1.14	2.50	1.38
Group 2	2.27	.82	2.32	.72	2.41	1.19	2.19	.94
Group 3	2.16	1.03	2.52	1.34	2.56	1.50	2.84	1.38
Group 4	2.00	.99	2.38	1.39	2.32	.90	2.17	.64
7. Adjective pair: "beneficial-harmful"								
Group 1	1.96	.94	2.22	.85	2.28	.79	2.25	.79
Group 2	2.35	.75	2.29	.60	2.19	.79	2.23	.95
Group 3	2.08	1.15	2.39	1.12	2.72	1.43	2.52	1.09
Group 4	1.96	.87	2.08	.94	2.72	1.24	2.29	1.12
8. Adjective pair: "successful-unsuccessful"								
Group 1	2.26	1.26	2.48	1.34	2.44	1.47	2.54	1.47
Group 2	2.85	1.32	2.43	.57	2.37	1.01	2.38	.90
Group 3	2.36	1.32	2.52	1.24	2.64	1.63	2.60	1.29
Group 4	2.08	1.22	2.32	1.28	2.52	1.09	2.21	.78
C. Concept: "School Administrators"								
1. Adjective pair: "good-bad"								
Group 1	1.41	.50	1.67	.68	1.68	.80	1.92	.93
Group 2	1.81	.69	1.61	.63	1.78	.75	2.07	1.07
Group 3	1.52	.82	1.74	.69	1.84	.94	1.96	.89
Group 4	1.77	1.11	1.96	1.00	2.04	1.00	1.96	.79
2. Adjective pair: "meaningless-meaningful"								
Group 1	1.48	.75	1.56	.64	1.60	.91	1.92	.93
Group 2	2.27	1.04	2.36	1.52	2.07	.87	2.19	1.10
Group 3	1.72	1.06	1.91	.67	1.88	.80	2.20	1.12
Group 4	1.73	.96	2.38	1.39	2.16	1.03	2.00	.89
3. Adjective pair: "fair-unfair"								
Group 1	1.41	.69	1.74	.71	1.80	.87	1.88	.95
Group 2	2.12	1.31	1.79	.83	1.96	1.19	2.15	1.22
Group 3	1.76	1.13	2.09	.90	1.92	.76	1.88	.83
Group 4	1.88	.91	2.19	1.20	2.24	1.09	2.04	.75
4. Adjective pair: "negative-positive"								
Group 1	1.52	1.22	1.56	.75	1.76	.78	1.63	.65
Group 2	1.81	.98	1.54	.88	1.85	.77	1.85	.83
Group 3	1.46	.51	1.96	1.15	1.92	1.08	1.80	.71
Group 4	1.88	.95	1.77	.71	2.12	.88	2.04	1.04

TABLE 28--Continued

Treatment Groups	Round 1		Round 2		Round 3		Round 4	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
5. Adjective pair: "unsuccessful-successful"								
Group 1	1.33	.62	1.93	1.14	1.60	.71	1.79	.88
Group 2	2.04	1.11	1.89	1.03	2.19	1.08	2.23	1.18
Group 3	1.60	.71	1.74	.69	2.24	1.33	1.92	.76
Group 4	1.77	.82	1.88	.91	2.04	.94	2.00	.98
6. Adjective pair: "valuable-worthless"								
Group 1	1.48	.51	1.89	.70	2.04	1.21	1.75	.74
Group 2	2.12	.77	1.89	.89	2.19	.96	1.92	.89
Group 3	1.92	.95	2.13	1.10	2.08	1.00	2.40	1.38
Group 4	2.27	1.00	2.04	.72	2.28	1.02	2.25	1.07
7. Adjective pair: "harmful-beneficial"								
Group 1	1.73	1.12	1.85	.99	1.92	1.14	2.04	1.04
Group 2	1.96	.77	1.85	.71	2.07	1.17	2.12	1.24
Group 3	2.08	1.12	2.22	1.17	2.36	1.15	2.28	1.02
Group 4	2.27	1.12	2.08	.98	2.20	.87	2.17	1.13
D. Concept: "Homework"								
1. Adjective pair: "good-bad"								
Group 1	2.70	1.17	3.04	1.16	2.96	1.14	2.79	1.14
Group 2	2.92	1.29	3.07	1.30	3.11	1.19	3.11	1.25
Group 3	3.16	1.43	3.13	1.39	3.08	1.35	3.16	1.11
Group 4	2.92	1.35	2.92	1.06	2.96	.94	2.92	1.26
2. Adjective pair: "meaningless-meaningful"								
Group 1	2.63	1.25	2.70	1.07	2.96	1.27	2.92	1.21
Group 2	2.88	1.18	3.00	1.44	2.85	1.26	2.92	1.20
Group 3	2.96	1.37	2.87	1.33	3.04	1.43	3.20	1.19
Group 4	2.65	1.36	2.85	1.08	2.92	1.19	2.83	1.17
3. Adjective pair: "fair-unfair"								
Group 1	3.08	1.26	2.38	.85	2.84	1.25	2.70	1.06
Group 2	3.15	1.64	2.93	1.33	3.07	1.52	2.85	1.41
Group 3	2.68	1.31	3.17	1.40	3.08	1.35	3.16	1.31
Group 4	2.77	1.42	3.00	1.47	2.88	1.15	2.79	1.38
4. Adjective pair: "negative-positive"								
Group 1	3.33	1.07	3.52	1.01	3.20	1.12	3.57	1.20
Group 2	3.96	1.37	3.67	1.36	3.96	1.32	3.58	1.27
Group 3	3.60	1.12	4.43	1.41	4.04	1.27	3.83	1.27
Group 4	3.85	1.32	3.96	1.37	4.08	1.19	3.54	1.35

TABLE 28--Continued

Treatment Groups	Round 1		Round 2		Round 3		Round 4	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
5. Adjective pair: "unpleasant-pleasant"								
Group 1	2.74	1.35	3.00	1.39	2.88	1.62	2.75	1.23
Group 2	3.04	1.59	3.04	1.80	3.00	1.39	2.77	1.53
Group 3	2.92	1.41	3.09	1.44	2.96	1.34	3.20	1.35
Group 4	3.28	1.65	3.04	1.46	2.92	1.50	3.25	1.48
6. Adjective pair: "necessary-unnecessary"								
Group 1	2.81	1.21	2.56	1.12	2.76	1.36	2.58	.93
Group 2	2.65	1.33	3.04	1.48	2.81	1.24	2.62	1.17
Group 3	2.60	1.08	2.78	1.24	3.04	1.33	2.96	1.17
Group 4	2.85	1.35	2.96	1.15	2.72	1.17	2.88	1.15
7. Adjective pair: "valuable-worthless"								
Group 1	2.48	.89	2.59	1.05	2.64	1.04	2.75	1.11
Group 2	2.62	1.20	2.93	1.27	2.70	1.17	2.54	.99
Group 3	2.60	.96	2.87	1.55	3.08	1.38	3.00	1.16
Group 4	2.58	1.03	2.88	1.18	2.84	1.31	2.88	1.12
8. Adjective pair: "harmful-beneficial"								
Group 1	2.96	1.28	2.85	1.12	3.00	1.16	3.04	1.27
Group 2	3.35	1.44	3.39	1.52	3.07	1.30	3.31	1.26
Group 3	3.04	1.14	3.39	1.53	3.16	1.11	3.56	1.29
Group 4	3.19	1.30	3.35	1.38	3.28	1.34	3.00	1.22
E. Concept: "Teaching Work Load"								
1. Adjective pair: "good-bad"								
Group 1	2.46	1.24	2.77	1.21	2.63	1.38	2.46	1.10
Group 2	3.36	1.75	3.54	1.84	3.41	1.76	3.23	1.77
Group 3	2.52	1.16	2.52	1.47	2.56	1.23	2.48	1.33
Group 4	2.96	1.64	3.46	1.77	3.36	1.55	2.80	1.22
2. Adjective pair: "fair-unfair"								
Group 1	2.69	1.38	2.58	1.10	2.25	1.07	2.25	.79
Group 2	2.60	1.26	2.85	1.77	2.81	1.57	3.27	1.78
Group 3	2.20	1.12	2.43	1.47	2.60	1.29	2.56	1.50
Group 4	2.65	1.16	3.31	1.57	3.28	1.43	2.63	1.35
3. Adjective pair: "interesting-boring"								
Group 1	1.46	.65	1.69	.74	1.63	.71	1.79	.88
Group 2	1.72	.79	2.00	1.14	2.22	1.42	2.00	1.20
Group 3	1.52	.65	1.70	.77	1.79	.59	1.67	.70
Group 4	1.73	.92	2.04	1.28	2.00	.76	2.04	1.08

TABLE 28 --Continued

Treatment Groups	Round 1		Round 2		Round 3		Round 4	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
4. Adjective pair: "unpleasant-pleasant"								
Group 1	2.04	.98	2.46	1.27	2.46	1.53	2.17	.96
Group 2	2.40	1.32	2.85	1.51	2.56	1.50	2.65	1.55
Group 3	1.96	.74	2.78	1.65	2.38	1.10	2.16	1.14
Group 4	2.19	1.30	2.50	1.33	2.20	1.04	2.33	1.20
5. Adjective pair: "necessary-unnecessary"								
Group 1	1.54	.65	1.85	1.01	1.79	.83	2.08	1.14
Group 2	2.44	1.76	2.07	1.11	2.04	1.19	2.62	1.65
Group 3	1.88	1.09	1.78	1.13	2.12	1.17	2.04	1.02
Group 4	1.77	.99	2.31	1.54	2.28	1.17	2.13	1.26
6. Adjective pair: "valuable-worthless"								
Group 1	1.77	.71	2.15	.93	1.92	.72	2.17	1.17
Group 2	2.16	1.03	2.37	1.12	2.48	1.28	2.38	1.33
Group 3	2.12	1.24	2.17	.89	1.84	.62	1.92	.95
Group 4	2.08	1.09	2.08	1.16	2.20	1.00	2.13	.85
7. Adjective pair: "harmful-beneficial"								
Group 1	1.96	1.00	2.58	1.45	2.08	1.06	2.00	1.06
Group 2	3.00	1.68	2.81	1.52	2.74	1.48	2.54	1.66
Group 3	2.24	1.51	2.35	1.34	2.12	1.24	1.92	1.12
Group 4	2.15	1.35	2.73	1.66	2.44	1.29	2.63	1.50
F. Concept: "Grading"								
1. Adjective pair: "good-bad"								
Group 1	3.30	1.73	3.19	1.30	3.16	1.41	3.25	1.48
Group 2	3.92	1.62	3.70	1.66	3.69	1.44	3.93	1.41
Group 3	3.20	1.63	3.30	1.80	3.08	1.61	3.24	1.56
Group 4	3.12	1.56	3.23	1.53	3.36	1.44	3.12	1.17
2. Adjective pair: "meaningless-meaningful"								
Group 1	2.74	1.13	3.04	1.22	2.68	1.22	3.25	1.48
Group 2	3.50	1.53	3.46	1.69	3.04	1.34	3.73	1.69
Group 3	3.20	1.61	3.39	1.85	2.96	1.57	3.08	1.47
Group 4	2.46	1.21	3.04	1.46	3.04	1.37	2.83	1.20
3. Adjective pair: "fair-unfair"								
Group 1	3.52	1.63	3.26	1.40	2.88	1.45	3.04	1.33
Group 2	4.12	1.58	3.70	1.59	3.81	1.60	3.54	1.36
Group 3	3.33	1.37	3.57	1.59	3.08	1.53	3.36	1.50
Group 4	3.35	1.33	3.35	1.41	3.36	1.60	3.00	1.25

TABLE 28--Continued

Treatment Groups	Round 1		Round 2		Round 3		Round 4	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
4. Adjective pair: "unpleasant-pleasant"								
Group 1	4.41	1.53	4.70	1.35	3.92	1.35	4.29	1.49
Group 2	5.46	1.27	5.39	1.34	4.96	1.48	4.77	1.45
Group 3	4.20	1.58	4.22	1.57	4.33	1.66	4.16	1.57
Group 4	3.96	1.64	4.60	1.53	4.63	1.50	4.46	1.44
5. Adjective pair: "necessary-unnecessary"								
Group 1	2.44	1.50	2.44	1.31	2.32	1.44	2.17	1.24
Group 2	2.77	1.63	2.36	1.42	2.67	1.47	2.38	1.39
Group 3	2.56	1.78	2.65	1.50	2.60	1.56	2.64	1.52
Group 4	2.08	1.09	2.08	.94	2.68	1.25	2.71	1.27
6. Adjective pair: "valuable-worthless"								
Group 1	2.93	1.44	3.11	1.25	2.68	.99	2.58	.65
Group 2	3.46	1.45	2.96	1.32	2.67	1.04	2.81	1.30
Group 3	2.83	1.44	2.78	1.00	2.72	1.21	2.92	1.44
Group 4	2.54	1.14	2.69	1.09	3.00	1.26	2.96	1.16
7. Adjective pair: "harmful-beneficial"								
Group 1	3.04	1.46	3.22	1.28	2.68	1.03	3.08	1.18
Group 2	3.77	1.37	4.00	1.68	3.33	1.36	3.62	1.27
Group 3	3.21	1.64	3.30	1.52	3.04	1.57	3.28	1.51
Group 4	3.04	1.46	2.96	1.18	3.40	1.41	3.42	1.38
G. Concept: "Teaching As a Career" 1. Adjective pair: "good-bad"								
Group 1	1.26	.53	1.37	.57	1.40	.50	1.38	.58
Group 2	1.19	.40	1.43	.74	1.56	1.09	1.52	.85
Group 3	1.28	.54	1.26	.45	1.40	.76	1.32	.63
Group 4	1.23	.52	1.27	.53	1.52	.71	1.44	.58
2. Adjective pair: "meaningless-meaningful"								
Group 1	1.11	.32	1.22	.42	1.48	1.05	1.71	1.63
Group 2	1.42	1.21	1.36	.62	1.41	1.05	1.35	.56
Group 3	1.36	1.22	1.43	1.08	1.44	1.23	1.44	1.08
Group 4	1.12	.33	1.31	.68	1.44	.51	1.58	.93
3. Adjective pair: "fair-unfair"								
Group 1	1.63	.74	1.58	.90	1.68	.95	1.79	.93
Group 2	2.31	1.26	2.25	1.21	2.15	1.32	2.27	1.25
Group 3	1.68	.85	1.77	1.07	1.64	1.00	1.96	1.24
Group 4	1.62	.90	1.96	1.34	1.80	1.00	1.88	.95

TABLE 28--Continued

Treatment Groups	Round 1		Round 2		Round 3		Round 4	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
4. Adjective pair: "unpleasant-pleasant"								
Group 1	1.41	.57	1.56	.64	1.36	.49	1.50	.78
Group 2	1.65	.75	1.75	1.11	1.70	1.07	1.50	.65
Group 3	1.36	.76	1.39	.50	1.36	.57	1.48	.87
Group 4	1.46	.58	1.31	.55	1.56	.71	1.50	.59
5. Adjective pair: "unsuccessful-successful"								
Group 1	1.07	.27	1.19	.40	1.32	.56	1.29	.62
Group 2	1.15	.37	1.32	.55	1.74	1.32	1.58	1.39
Group 3	1.12	.33	1.26	.45	1.36	.57	1.24	.52
Group 4	1.27	.60	1.31	.69	1.48	.71	1.54	.83
6. Adjective pair: "valuable-worthless"								
Group 1	1.67	.73	1.52	.51	1.64	.70	1.96	1.33
Group 2	2.00	.85	2.25	1.43	2.07	1.24	1.88	.77
Group 3	1.52	.59	1.61	.66	1.58	.65	1.68	.80
Group 4	1.92	1.23	1.46	.71	1.67	.70	1.63	.71